USING RF-DNA FINGERPRINTS TO DISCRIMINATE ZIGBEE DEVICES IN AN OPERATIONAL ENVIRONMENT

THESIS

Clay K. Dubendorfer, Civilian, USAF

AFIT-ENG-13-M-15

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the United States Government.

AFIT-ENG-13-M-15

USING RF-DNA FINGERPRINTS TO DISCRIMINATE ZIGBEE DEVICES IN AN OPERATIONAL ENVIRONMENT

THESIS

Presented to the Faculty
Department of Electrical and Computer Engineering
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Electrical Engineering

Clay K. Dubendorfer, B.S.E.E.
Civilian, USAF

March 2013

AFIT-ENG-13-M-15

USING RF-DNA FINGERPRINTS TO DISCRIMINATE ZIGBEE DEVICES IN AN OPERATIONAL ENVIRONMENT

Clay K. Dubendorfer, B.S.E.E.
Civilian, USAF

Approved:

Michael A. Temple, PhD (Chairman) — 12 Mar 2013 Date

Lt. Col. Jeffrey D. Clark, PhD (Member) — 8 Mar 2013 Date

Robert F. Mills, PhD (Member) — 7 MAR 2013 Date

AFIT-ENG-13-M-15

Abstract

This research was performed to expand AFIT's Radio Frequency "Distinct Native Attribute" (RF-DNA) fingerprinting process to support IEEE 802.15.4 ZigBee communication network applications. Current ZigBee bit-level security measures include use of network keys and Media Access Control (MAC) lists which can be subverted through interception and spoofing using open-source hacking tools. This work addresses device discrimination using Physical (PHY) waveform alternatives to augment existing bit-level security mechanisms. ZigBee network vulnerability to outsider threats was assessed using Receiver Operating Characteristic (ROC) curves to characterize both *Authorized Device ID Verification* performance (granting network access to authorized users presenting *true* bit-level credentials) and *Rogue Device Rejection* performance (denying network access to unauthorized rogue devices presenting *false* bit-level credentials).

Radio Frequency 'Distinct Native Attribute' (RF-DNA) features are extracted from time-domain waveform responses of 2.4 GHz CC2420 ZigBee transceivers to enable human-like device discrimination. The fingerprints were constructed using a "hybrid" pool of emissions collected under a range of conditions, including anechoic chamber and an indoor office environment where dynamic multi-path and signal degradation factors were present. The RF-DNA fingerprints were input to a Multiple Discriminant Analysis, Maximum Likelihood (MDA/ML) discrimination process and a 1 vs. many "Looks most like?" classification assessment made. The hybrid MDA model was also used for 1 vs. 1 "Looks how much like?" verification assessment. ZigBee *Device Classification* performance was assessed using both full and reduced dimensional fingerprint sets. Reduced dimensional subsets were selected using Dimensional Reduction Analysis (DRA) by rank ordering 1) pre-classification Kolmogorov-Smirnov (KS)-Test p-values and 2) post-classification Generalized Relevance Learning Vector Quantization-

Improved (GRLVQI) λ_i feature relevance values. Assessment of Zigbee device ID verification capability included both *Authorized Device ID Verification* and *Rogue Device Rejection*.

Device Classification performance using full-dimensional fingerprints comprised of N_F=729 features achieved an arbitrary benchmark of average correct classification $\%C$>90% (across all devices) for $SNR \geq 10.0$ dB. Performance using DRA≈66% ($N_{\widehat{F}}$=243) reduced dimensional subsets was marginally poorer and yielded a "gain" of $G \approx -1.0$ dB at $\%C$=90% relative to full-dimensional performance; gain is the reduction in required SNR for two systems, methods, etc., to achieve a given $\%C$. Additional KS-Test and GRLVQI DRA feature selection was performed and classification performance assessed using the top-ranked $N_{\widehat{F}}$=200, 100, 50, and 25 features. Relative to the %C>90% benchmark, the KS-Test and GRLVQI selected feature sets required the same $SNR \approx 10.0$ dB ($N_{\widehat{F}}$=243) to $SNR \approx 18.0$ dB ($N_{\widehat{F}}$=50). For $N_{\widehat{F}}$=25, KS-Test selected features failed to meet the benchmark while GRLVQI selected features achieved the benchmark at $SNR \approx 30.0$ dB.

Authorized Device ID Verification performance was evaluated using the $N_{\widehat{F}}$=50 DRA feature set. Results indicate the existence of a device dependent threshold whereby *all* authorized devices achieve an arbitrary True Verification Rate (TVR>90%) and False Verification Rate (FVR<10%) benchmark for both DRA methods. *Rogue Device Rejection* was assessed using unauthorized rogue devices, with each rogue device falsely presenting a claimed ID matching each of the authorized device IDs. Considering an arbitrary Rogue Rejection Rate (RRR>90%) benchmark, ROC curve analysis for *Rogue Device Rejection* indicated that performance using KS-Test and GRLVQI selected feature sets were consistent. The KS-test DRA selected feature sets achieved RRR>90% in 21, 29, and 30 of 36 rogue scenarios using $N_{\widehat{F}}$=100, 50, and 25 top-ranked features, respectively. Similarly, the GRLVQI DRA selected features achieved RRR>90% in 23, 28, and 30 of the 36 rogue scenarios using $N_{\widehat{F}}$=100, 50, and 25 top-ranked features, respectively.

Acknowledgments

I would like to thank Dr. Temple for the continuous guidance and mentoring during my research efforts. The abundant time, effort, and positive attitude was greatly appreciated and helped make this a rewarding and worthwhile experience. A special thanks goes out to fellow classmates that were involved and assisted in the research. I would also like to thank my friends and family for the continuing support and perspective provided during the process of this research.

Clay K. Dubendorfer

Table of Contents

List of Figures

List of Tables

Table Page

List of Acronyms

Acronym	Definition
ADC	Analog-to-Digital Converter
AFIT	Air Force Institute of Technology
APL	Application
AWGN	Additive White Gaussian Noise
CDF	Cumulative Distribution Functions
DRA	Dimensional Reduction Analysis
GRLVQI	Generalized Relevance Learning Vector Quantization-Improved
GSM	Global System for Mobile Communication
IEEE	Institute of Electrical and Electronics Engineers
IF	Intermediate Frequency
I-Q	In-phase and Quadrature
KS	Kolmogorov-Smirnov
MAC	Media Access Control
MDA	Multiple Discriminant Analysis
MDA/ML	Multiple Discriminate Analysis, Maximum Likelihood
ML	Maximum Likelihood
MVG	Multivariate Gaussian
NWK	Network
OSI	Open Systems Interconnect
PHR	PHY Header
PHY	Physical
PPDU	PHY Protocol Data Unit
PSD	Power Spectral Density

Acronym	Definition
RAR	Rogue Accept Rate
RF	Radio Frequency
ROC	Receiver Operating Characteristics
ROI	Region Of Interest
RRR	Rogue Reject Rate
SFD	Start-of-Frame Delimiter
SHR	Synchronization Header
SNR	Signal to Noise Ratio
TI	Texas Instruments
TVR	True Verification Rate
WPAN	Wireless Personal Area Networks

USING RF-DNA FINGERPRINTS TO DISCRIMINATE ZIGBEE DEVICES IN AN OPERATIONAL ENVIRONMENT

I. Introduction

1.1 Operational Motivation

Wireless Personal Area networks (WPANs) are increasing in popularity and are widely deployed in office buildings, factories, home networks, and hospitals. The Institute of Electrical and Electronics Engineers (IEEE) 802.15.4 Media Access Control (MAC) and Physical-layer (PHY) standards provide a low power, low-data-rate WPAN foundation on which network (NWK) and application (APL) layers are built, such as the ZigBee specification [26]. ZigBee networks' low implementation costs and low-complexity make them a viable solution for applications such as industrial control and monitoring [14], home automation, remote metering [46], patient vital sign monitoring [7], security systems [45], and asset tracking [50]. Depending on the application, ZigBee networks transmit sensitive personal information, control physical systems (valves, fans, lighting, doors, etc.), and monitor critical sensors. Improved security measures is an essential component in allowing ZigBee-based networks to be highly reliable and secure. The need for improving network security is motivated by open source tools such as KillerBee [49] and Api-do [41] which increase ZigBee network vulnerability and enable unauthorized rogue devices to conduct packet replay, network key sniffing, MAC address spoofing, malicious network impersonation, and denial of service type attacks.

Wireless networks are characterized by the seven layer Open Systems Interconnect (OSI) model such as shown in Fig. 1.1 [1]. Traditionally, systems have predominantly relied on ”bit-level” security mechanisms implemented in the Network (NWK) and

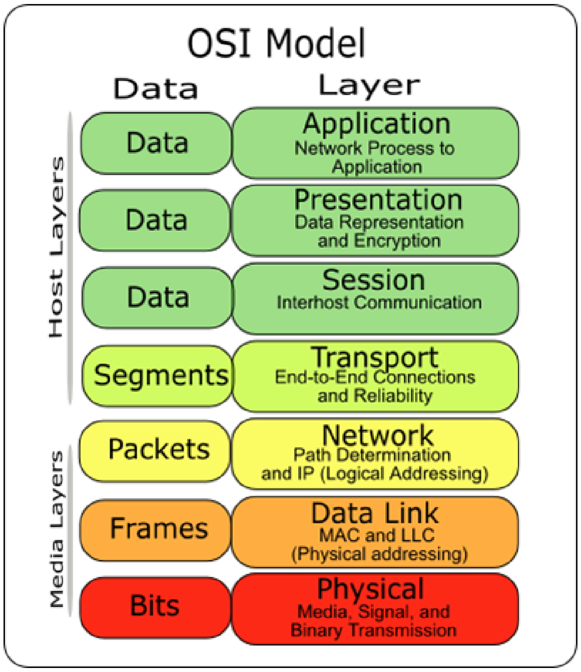

OSI Model
Data
Layer
Host Layers
Media Layers
Data
Application
Network Process to Application
Data
Presentation
Data Representation and Encryption
Data
Session
Interhost Communication
Segments
Transport
End-to-End Connections and Reliability
Packets
Network
Path Determination and IP (Logical Addressing)
Frames
Data Link
MAC and LLC (Physical addressing)
Bits
Physical
Media, Signal, and Binary Transmission

Figure 1.1: Multi-layer Open Systems Interconnect (OSI) network model [1].

Data Link (DLL) layers while generally ignoring the potential for PHY-layer security augmentation. Exploiting this potential has been a major motivation for ongoing research at Air Force Institute of Technology (AFIT) which exploits wireless device PHY waveform features. This is accomplished using Radio Frequency ‘Distinct Native Attribute’ (RF-DNA) fingerprints which provide unique, human-like device discrimination using RF-DNA features that vary due to component manufacturing differences, component tolerances, design differences, and device aging. The inherent RF-DNA is difficult to mimic and replicate, allowing it to be useful in discrimination between multiple devices. PHY layer security using RF-DNA fingerprints is a viable solution for augmenting higher layer (NWK and DLL) bit-level security mechanisms.

1.2 Technical Motivation

AFIT's RF-DNA fingerprinting process has evolved into the process shown in Fig. 1.2. This process is constantly expanding by considering new signal types, new feature types, new classification methods, and new device ID verification methods. Over the past several years, extensive research has been conducted at AFIT [21, 23, 24, 28–30, 34, 35, 37–40, 42, 47, 48] and contribution has been made to a larger body of research being conducted by numerous researchers [8–10, 16–19, 27]. AFIT's research activity has predominately focused on RF-DNA fingerprinting for *Device Classification* using various wireless communication signal types, such as Global System for Mobile Communication (GSM) cellular phones [40, 47], IEEE 802.11 WiFi [21, 23, 24, 28, 29, 35, 42], and IEEE 802.16 WiMAX [34, 35, 37, 38, 48]. This research is no exception and the RF-DNA process is adopted here to assess IEEE 802.15.4 ZigBee *Device Classification*. However, there has been a recent shift in AFIT research and this research is among the first few efforts to consider *Device ID Verification* using RF-DNA fingerprints.

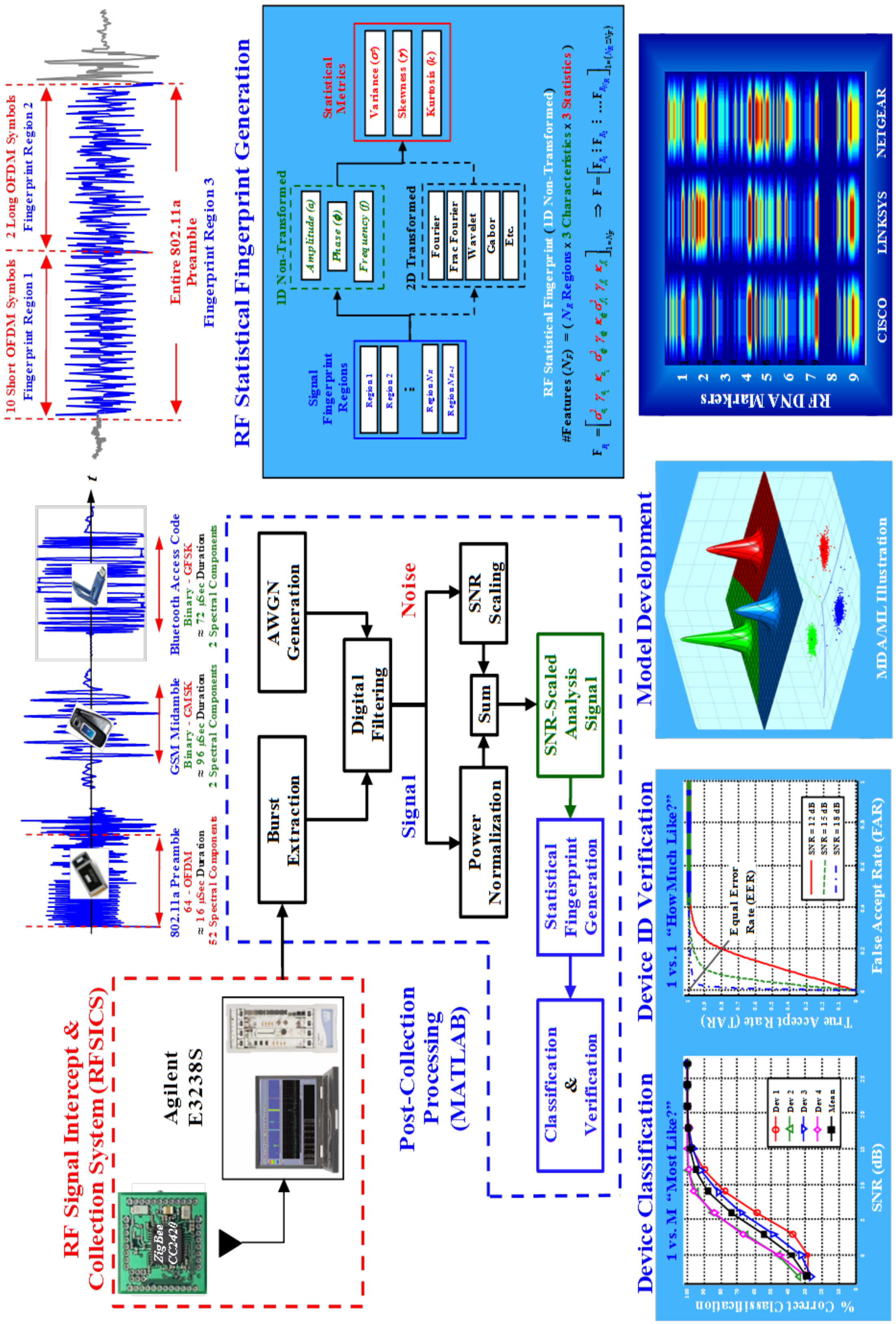


Figure 1.2: AFITs RF-DNA Fingerprinting Overview

1.3 Previous vs. Current Research

Table 1.1 provides a summary of technical areas that have been previously addressed and areas addressed under this research.

Table 1.1: *Technical Areas* in *Previous* related work and *Current* research contributions. The × symbol denots areas addressed.

Technical Area	Previous Work		Current Research	
	Addressed	Ref #	Addressed	Ref #
1D Time Domain (TD)	×	[8, 17, 28, 29, 43, 47] [42, 43, 47, 48]	×	[11, 12]
1D Spectral Domain (SD)	×	[38, 48]		
2D Wavelet Domain (WD)	×	[28–30]		
2D Gabor (GT/GWT)	×	[21, 34, 35, 37, 38]		
Signal Type				
802.11a WiFi	×	[21, 28–30, 35, 48]		
GSM Cellular	×	[39, 40, 47]		
802.16e WiMax	×	[34, 35, 38, 48]		
802.15.4 ZigBee	×	[31]	×	[11, 12]
Classifier Type				
MDA/ML	×	[28–30, 42, 43, 47, 48] [21, 31, 34, 38–40]	×	[11, 12]
GRLVQI	×	[21, 28, 29, 35, 37]		
LFS	×	[4–6, 21–24]		
Dimensional Reduction Analysis (DRA)				
GRLVQI	×	[28, 29, 33, 35, 37]	×	[12]
LFS	×	[20, 21]		
KS-Test	×	[31]	×	[12]
Device ID Verification				
Authorized Device	×	[35, 37]	×	[12]
Rogue Device Rejection	×	[35, 37]	×	[12]

1.4 Document Organization

The remainder of this document is organized as follows:

- **Chapter 2 - Background**: Provides fundamental information on ZigBee IEEE 802.15.4 signal structure. Describes the previously established procedure for extracting time-domain features. Explains Multiple Discriminant Analysis (MDA) model development and Maximum Likelihood (ML) Classification.

- **Chapter 3 - Research Methodology**: Describes the specific methodology used in this research to implement RF-DNA fingerprinting using experimentally collected ZigBee emissions, including emission collection and post-collection processing. Describes RF-DNA fingerprint quantitative Dimensional Reduction Analysis (DRA) methods, including: 1) pre-classification KS-Test p-value ranking, and 2) post-classification GRLVQI λ_i relevance ranking. Details the methodology used to perform ZigBee device discrimination, including *Device Classification*, *Authorized Device ID Verification*, and *Rogue Device Rejection*.

- **Chapter 4 - Results and Analysis**: Provides results and performance analysis for full-dimensional and DRA reduced dimensional RF-DNA fingerprinting using KS-Test and GRLVQI selected feature sets. *Device classification* performance for full-dimensional and reduced dimensional feature sets. This includes assessment of *Device Classification*, *Authorized Device ID Verification*, and *Rogue Device Rejection* capability.

- **Chapter 5 - Summary and Conclusions**: Presents a summary of research activity, significant results, and recommendations for future research.

II. Background

This chapter provides the technical background information supporting development of the methodology described in Chap. 3 and interpretation of results presented in Chap. 4. Section 2.1 provides details for ZigBee-based networks built on IEEE 802.15.4 standard for wireless low-data-rate Wireless Personal Area Networks (WPAN). Section 2.2 explains the process for generating RF-DNA fingerprints that are comprised of statistical features extracted from time-domain emission responses. A description of Multiple Discriminant Analysis (MDA) model development and Maximum Likelihood (ML) classification processes are described in Sections 2.3 and 2.4, respectively, and are the foundation for MDA/ML processing used in developing Chap. 3 methodology.

2.1 ZigBee Signal Structure

ZigBee technology is used for WPANs and is seen in many applications requiring a low data rate, long battery life, and low cost solution. These applications include home automation, industrial control and monitoring, remote sensing/metering, medical equipment and patient monitoring, asset tracking systems, security systems, lighting and temperature control, etc. ZigBee-based networks are built on the WPAN IEEE 802.15.4 standard which defines the Physical (PHY) and Media Access Control (MAC) layer structure. The ZigBee specification [51] defines the Network (NWK) layer specifications and provides a framework for application programming in the Application (APL) layer.

Figure 2.1 shows the MAC frame format and PHY layer structure used by ZigBee [26]. As described in the 2.4 GHz IEEE 802.15.4 standard, the PHY Protocol Data Unit (PPDU) packet structure consists of 1) a Synchronization Header (SHR) response which allows a receiving device to synchronize and lock onto the bit stream, 2) a PHY Header (PHR) response which contains frame length information, and 3) a variable length payload which carries the MAC sublayer frame. The SHR region in Fig. 2.2 is comprised of a 32-bit preamble and an 8-bit Start-of-Frame Delimiter (SFD) sequence. The preamble sequence is designed for acquisition of symbol chip timing and is composed of a 32-bit binary zero string. The SFD region is used to signify the end of preamble and consists of a predefined 8-bit sequence of [1 1 1 0 0 1 0 1]. Information contained with the SHR region remains constant and is independent of device emissions, individual device types, device applications, etc. Early research reported in [11] exploited the preamble-only region of ZigBee emissions for RF-DNA fingerprinting. Subsequent analysis revealed a greater level of device discrimination can be realized using the entire SHR region (preamble and SFD). Thus, the methodology described in Chap. 3 and results in Chap. 4 are based exclusively on RF-DNA exracted from the SHR region.

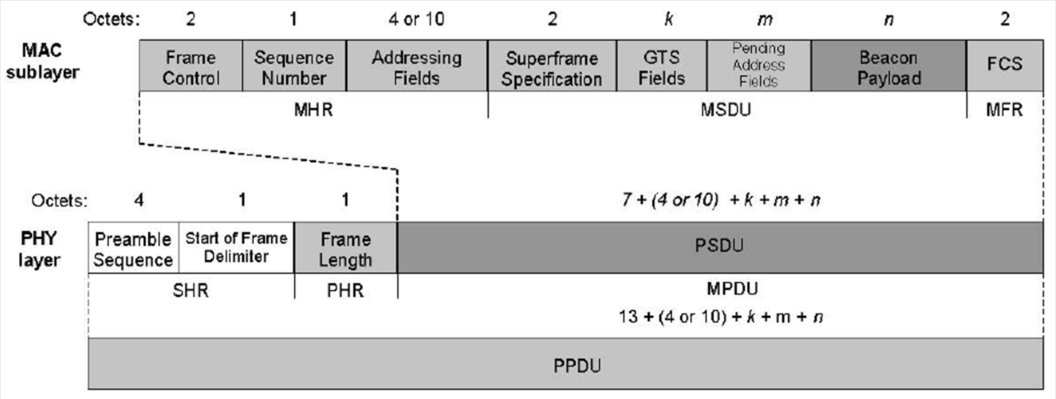
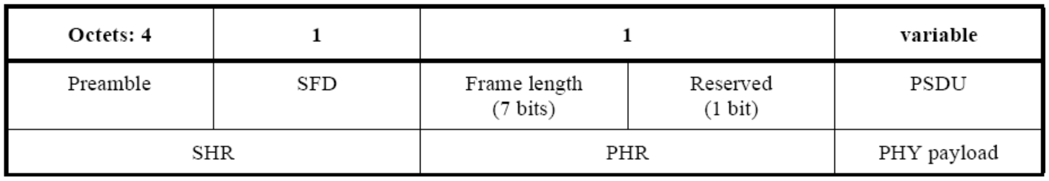


Figure 2.1: Data frame PHY and MAC layer structures for a ZigBee packet [26].

2.2 Time-Domain RF-DNA Fingerprint Generation

The RF-DNA fingerprints for an emission Time Domain (TD) response are derived from its instantaneous amplitude (a), phase (ϕ) and frequency (f) responses, as described in [11, 12, 30, 33, 39, 40, 43, 48]. The corresponding characteristic sequences, having



Octets: 4	1	1		variable
Preamble	SFD	Frame length (7 bits)	Reserved (1 bit)	PSDU
SHR		PHR		PHY payload

Figure 2.2: PHY Protocol Data Unit (PPDU) packet structure for IEEE 802.15.4 [26].

elements denoted by $a[n]$, $\phi[n]$, and $f[n]$, are generated using N_S complex I-Q signal samples $s[n]{=}s_I[n]{+}js_Q[n]$ from the specific Region Of Interest (ROI) in the collected signal where the mean value is removed (centered) and then normalized (division by maximum value) [30, 43]. Elements of the emission TD response are calculated by,

$$a[n] = \sqrt{s_I^2[n] + s_Q^2[n]}, \tag{2.1}$$

$$\phi[n] = \tan^{-1}\left[\frac{s_Q[n]}{s_I[n]}\right], \text{for } s_I[n] \neq 0, \tag{2.2}$$

$$f(n) = \frac{1}{2\pi}\left[\frac{d\phi(n)}{dt}\right]. \tag{2.3}$$

Mean removal and normalization for each of the N_S elements in characteristic sequences, $\{a[n]\}$, $\{\phi[n]\}$, and $\{f[n]\}$, is achieved using,

$$\bar{a}_c(n) = \frac{a[n] - \mu_a}{\max\limits_{n}\{a_c[n]\}}, \tag{2.4}$$

$$\bar{\phi}_c[n] = \frac{\phi[n] - \mu_\phi}{\max\limits_{n}\{\phi_c[n]\}}, \tag{2.5}$$

$$\bar{f}_c[n] = \frac{f[n] - \mu_f}{\max\limits_{n}\{f_c[n]\}}, \tag{2.6}$$

where $n = 1, 2, 3, \ldots, N_S$, and μ_a, μ_ϕ and μ_f are the means of $\{a[n]\}$, $\{\phi[n]\}$, and $\{f[n]\}$ calculated across N_S samples, and $\max\{\cdot\}$ denotes the maximum value of each feature sequence's centered magnitude.

RF-DNA fingerprints are compromised of statistical features extracted from instantaneous TD responses over a specific ROI in the collected signal [11, 12, 30, 33, 39, 40, 43,

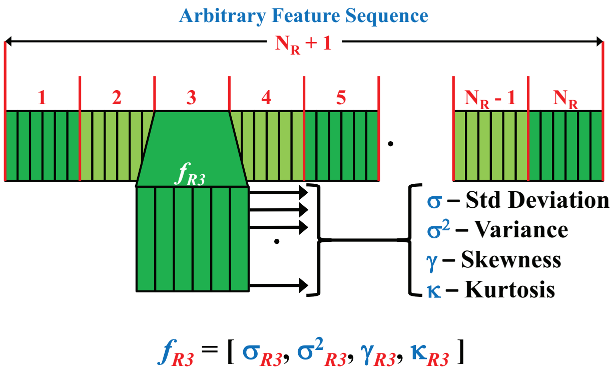


Figure 2.3: Representative illustration of regional fingerprint marker generation for an arbitrary ROI sequence using N_R+1 total subregions and N_M=4 statistical metrics [33].

48]. The selected ROI is a response region that is 1) ideally consistent across all collected signals, and 2) independent of data modulation and device ID information. As shown in Fig. 2.3, statistical RF-DNA features of standard deviation (σ), variance (σ^2), skewness (γ), and kurtosis (κ) are calculated over the ROI to form *regional fingerprint markers* generated by: 1) dividing each selected characteristic sequence $\{a[n]\}$,$\{\phi[n]\}$, and $\{f[n]\}$ into N_R contiguous, equal length subsequences such that N_S/N_R is an integer, 2) calculating N_M metrics for each subsequence, plus the entire fingerprinted region as a whole (N_R+1 total regions), and 3) arranging the metrics in a vector of the form,

$$F_{R_i} = [\sigma_{R_i}\ \sigma^2_{R_i}\ \gamma_{R_i}\ \kappa_{R_i}]_{1\times 4}\ , \tag{2.7}$$

where $i = 1, 2, \ldots, N_R + 1$. The N_M metrics for each subsequence are calculated from,

$$\mu = \frac{1}{N}\sum_{n=1}^{N} x[n]\ , \tag{2.8}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(x[n]-\mu)^2}\ , \tag{2.9}$$

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N}(x[n]-\mu)^2\ , \tag{2.10}$$

$$\gamma = \frac{1}{N\sigma^3} \sum_{n=1}^{N} (x[n] - \mu)^3 \ , \tag{2.11}$$

$$\kappa = \frac{1}{N\sigma^4} \sum_{n=1}^{N} (x[n] - \mu)^4 \ , \tag{2.12}$$

where $x[n]$ is the n^{th} feature vector element and N is the total number of samples in each subsequence used to calculate the statistic.

The marker vectors from (2.7) are concatenated to form the *composite characteristic vector* for each characteristic and are given by,

$$\mathbf{F} = [F_{R_1} \vdots F_{R_2} \vdots F_{R_3} \ldots F_{R_{N_{R+1}}}]_{1\times[N_M\times(N_R+1)]} \tag{2.13}$$

If only one signal characteristic is used (a, ϕ, or f), the expression in (2.13) represents the final classification fingerprint. When all $N_C = 3$ signal characteristics are used, the final RF fingerprint is generated by concatenating vectors from (2.13) according to

$$\mathbf{F} = [\mathbf{F}^a \vdots \mathbf{F}^\phi \vdots \mathbf{F}^f]_{1\times[N_M\times(N_R+1)\times N_C]} \tag{2.14}$$

The final full-dimensional RF fingerprint (2.14) is a vector comprised of N_F features, where

$$N_F = N_M \times (N_R + 1) \times N_C \tag{2.15}$$

2.3 Multiple Discriminant Analysis (MDA)

The research methodology presented in Chap. 3 is based on fundamental MDA concepts described in this section. MDA is a linear method of projecting high-dimensional data into a lower-dimensional space that best separates data in a least-squares sense [13]. MDA is performed on RF-DNA fingerprints to reduce the feature dimensionality and aid in the development of a class (device) specific model as described in 3.5.1.

MDA is an extension of Fisher's Linear Discriminant process when discrimination of two or more classes is required (N_C>2). MDA reduces input feature dimensionality by projecting N_F-dimensional input features into a (N_C−1)-dimensional subspace, where it is assumed that $N_F \geq N_C$. This linear transformation (projection) is performed with a goal toward maximizing the out-of-class separation (class mean differences) and minimizing within-class spread (variance within each class) of input data projections [13].

The out-of-class (inter-class, $\mathbf{S}_b$) and within-class (intra-class, $\mathbf{S}_w$) scatter matrices in MDA are computed as [44],

$$\mathbf{S}_b = \sum_{i=1}^{N_C} P_i \boldsymbol{\Sigma}_i \ , \tag{2.16}$$

$$\mathbf{S}_w = \sum_{i=1}^{N_C} P_i (\mu_i - \mu_0)(\mu_i - \mu_0)^T \ , \tag{2.17}$$

with class covariance ($\boldsymbol{\Sigma}_i$) and global mean (μ_0) calculated as follows,

$$\Sigma_i = E[(x - \mu_i)(x - \mu_i)^T] \ , \tag{2.18}$$

$$\mu_0 = \sum_{i=1}^{N_C} P_i \mu_i \ , \tag{2.19}$$

where μ_i is the mean and P_i is the prior probability of each N_C class. The within-class scatter matrix in (2.17) provides a measure of probability-weighted class feature variance and the out-of-class scatter matrix in (2.16) provides a measure of the average (over all classes) distance between individual class means from the respective global mean.

The N_F-dimensional input RF-DNA fingerprint vectors, $\mathbf{F}$ from (2.13), are projected into the lower (N_C-1)-dimensional subspace using,

$$\hat{\mathbf{f}} = \mathbf{W}^T \mathbf{F} \ , \tag{2.20}$$

where $\mathbf{W}$ is the $N_F \times (N_C-1)$ transformation (projection) matrix formed from the N_C-1 eigenvectors of $\mathbf{S}_w^{-1}\mathbf{S}_b$ and $\hat{\mathbf{f}}$ is the projected RF-DNA fingerprint. This linear projection by matrix $\mathbf{W}$ results in the optimal ratio between inter-class distances and intra-class variances [44]. Figure 2.4 shows two possible representative MDA projection transformations ($\mathbf{W}_1$ and $\mathbf{W}_2$) for N_C=3 classes onto a 2-dimensional subspace; for this illustration $\mathbf{W}_1$ provides the "best" class separation.

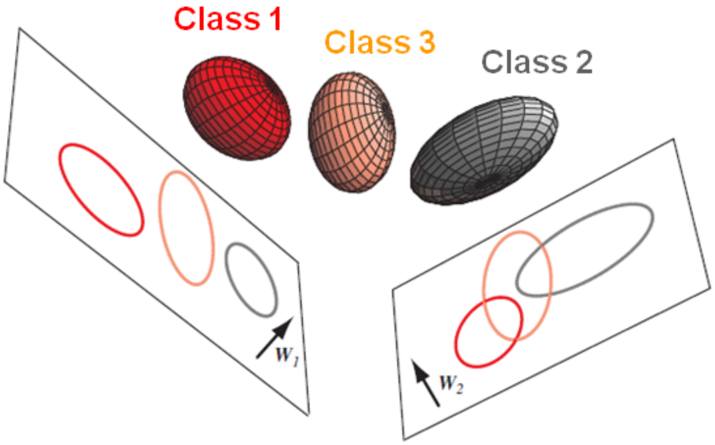


Figure 2.4: Representative projections for N_C=3 classes projected onto 2-dimensional subspaces using $\mathbf{W}_1$ and $\mathbf{W}_2$ [13]; $\mathbf{W}_1$ is more optimal in this case.

2.4 Maximum Likelihood (ML) Classification

This section describes the ML classification process used in the research methodology described in Chapter 3. When considering N_C>2 classes comprised of N_F-dimensional input features, ML classification can be performed using an MDA-based model described in Sect. 2.3; the "model" consists of projection matrix $\mathbf{W}$. The available input data set for each of the N_C classes is divided into *Training* and *Testing* data sets, with the *Training* set used for MDA model development per Sect. 2.3 and *Testing* set used for ML classification.

For ML classification, the MDA model ($\mathbf{W}$) is first used to project the *Training* set for all N_C classes into the Fisher space. Class specific projected means ($\hat{\mu}_i$) and covariances ($\hat{\Sigma}_i$) are then computed for i=1, 2, . . . , N_C. The projected data is assumed to be multivariate Gaussian distributed with class-dependent means of $\hat{\mu}_i$ and class-dependent covariances of

$\hat{\Sigma}_i$. Alternately, identical covariances can be assumed and a pooled covariance estimate $\hat{\Sigma}_P$ used for all classes:

$$\hat{\Sigma}_P = \frac{1}{N_C} \sum_{i=1}^{N_C} \hat{\Sigma}_i \ . \tag{2.21}$$

The assumed MVG distributions effectively represent posterior conditional probabilities that can be used to measure class likelihood for projected *Testing* fingerprint $\hat{\mathbf{f}}$. For a pooled covariance estimate, likelihood estimation can be implemented as [33, 44],

$$P\left(\hat{\mathbf{f}}|N_{Ci}\right) = \frac{1}{(2\pi)^{(N_C-1)/2} \det\left(\hat{\Sigma}_P\right)^{1/2}} \cdot \exp(\mathcal{F}_e) \ , \tag{2.22}$$

where,

$$\mathcal{F}_e = -\frac{1}{2}\left(\hat{\mathbf{f}} - \hat{\mu}_i\right)^T \left(\hat{\mathbf{\Sigma}}_P\right)^{-1} \left(\hat{\mathbf{f}} - \hat{\mu}_i\right) \ . \tag{2.23}$$

Class likelihood values are used for ML classification based on Bayesian decision theory by assigning a class label to subsequent *Testing* data. In the case of N_C classes, a given projected *Testing* fingerprint $\hat{\mathbf{f}}$ is assigned to class c_i according to,

$$P\left(c_i|\hat{\mathbf{f}}\right) > P\left(c_j|\hat{\mathbf{f}}\right) \quad \forall j \neq i \ , \tag{2.24}$$

where i=$1, 2, \ldots, N_C$ and $P\left(c_i|\hat{\mathbf{f}}\right)$ is the conditional posterior probability that $\hat{\mathbf{f}}$ belongs to class c_i. The conditional posterior probability $P\left(c_i|\hat{\mathbf{f}}\right)$ is found by applying Bayes' Rule and using class likelihood values as shown [33, 44]:

$$P\left(c_i|\hat{\mathbf{f}}\right) = \frac{P\left(\hat{\mathbf{f}}|c_i\right) P(c_i)}{P\left(\hat{\mathbf{f}}\right)} \tag{2.25}$$

where prior probabilities are assumed equal for all classes ($P(c_i)$=$1/N_C$) and thus can be neglected when making (2.24) comparison. Since (2.25) is applied for a given projected $\hat{\mathbf{f}}$ fingerprint, $P\left(\hat{\mathbf{f}}\right)$ remains constant across all c_i and can also be neglected as well. Using the decision criteria from (2.24), projected *"testing"* fingerprints $\hat{\mathbf{f}}$ are assigned a class label c_i based on maximum posterior probability, with correct classification occurring when the assigned class label matches the true class label. This ML classification process is used in the research methodology to perform device classification as described 3.5.2.

III. Research Methodology

This chapter provides the methodology used to conduct this research and obtain results presented in Chap. 4. Topics are presented sequentially relative to the RF-DNA processing overview shown in Fig. 3.1. This process begins with ZigBee device signal collections made in three different environment scenarios as described in Section 3.1. Section 3.2 explains the post-processing procedure that is performed on collected emissions prior to RF-DNA fingerprint generation. Section 3.3 provides specifics on how ZigBee time-domain features are used to generate RF-DNA fingerprints. Dimensional Reduction Analysis (DRA) and two *quantitative* selection methods 1) pre-classification Kolmogorov-Smirnov (KS)-Test p-value ranking and 2) post-classification Generalized Relevance Learning Vector Quantization-Improved (GRLVQI) λ_i relevance ranking are introduced in Section 3.4. As explained in Section 3.5, RF-DNA fingerprints were input to a Multiple Discriminant Analysis (MDA) process and the resultant model used for both Maximum Likelihood (ML) *Device Classification* (Section 3.5.2) and device ID verification, specifically *Authorized Device ID Verification* (Section 3.5.3.1) and *Rogue Device Rejection* (Section 3.5.3.2).

3.1 Signal Collection

An Agilent E3238S [2] receiver (Rx) was used to collect emissions from ten Texas Instruments (TI) CC2420 2.4 GHz IEEE 802.15.4 ZigBee devices (denoted herein as Dev1, Dev2, ..., Dev10). The Agilent Rx can collect signals at an Radio Frequency (RF) center frequency spanning f_c=20.0 MHz to f_c=6.0 GHz using a tunable RF filter with an instantaneous bandwidth of W_{RF}=36.0 MHz. The selected frequency band is down-converted to an Intermediate Frequency (IF) of f_{IF}=70 MHz and digitized by an N_b=12 bit Analog-to-Digital Converter (ADC) operating at a sampling rate of f_s=95 Mega-Samples-

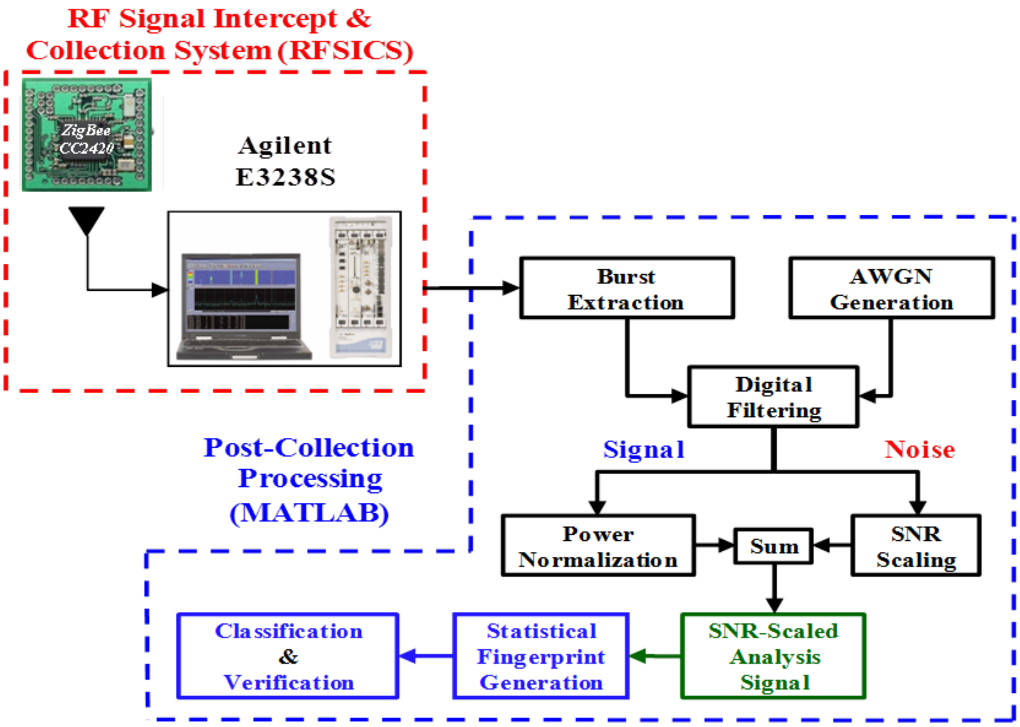


Figure 3.1: Overview of AFIT's RF-DNA Fingerprinting Process [36].

per-second (MSps), digitally down-converted to near baseband, baseband filtered with a specific (user defined) bandwidth W_{BB}, and automatically sub-sampled at a rate based on (W_{BB}) in accordance with Nyquist criteria requirements. All resultant collected samples are stored as complex In-phase and Quadrature (I-Q) data in a *.cap* file format [3].

Prior to device signal collections all CC2420 radio transceivers were programmed to transmit 2.4 GHz IEEE 802.15.4 compliant packets (bursts, pulses, etc.) with an arbitrary payload at a rate of 14 transmissions-per-second. The arbitrary payload is irrelevant to this research because RF-DNA fingerprints are generated from the Synchronization Header (SHR) region within the transmitted bursts. For each transmitting (Tx) CC2420 device, a total of N_B=1000 burst responses were collected under three operating conditions, including: 1) both the Tx and RX antenna inside a Ramsey STE3000B RF shielded anechoic chamber (“CAGE”) as done in [11, 31], 2) the Tx and Rx having a clear Line-of-Sight (“LOS”) path down a hallway–location A in Fig. 3.2 [12]) , and 3) the Tx and Rx on opposite sides of a wall (“WALL”)–location B shown in Fig. 3.2 [12].

During “Cage” collections the Tx position was consistently maintained at 20 cm from a dipole antenna in an RF-absorbent Ramsey STE3000B test enclosure that was connected to the Agilent Rx input by a shielded cable. For the experimental “LOS” collections (location A), the devices under test (Tx) were placed 5.0 m from a stationary 6 dB gain Ramsey LPY2 log periodic antenna [32] attached to the Rx. For “WALL” collections (location B), the devices (Tx) were placed behind an interior wall (5.5 m from Rx) consisting of 1.6 cm-thick drywall separated by 9.2 cm steel studs spaced 40.6 cm on center, for a total thickness of 12.4 cm, where fiberglass sound batting fills inter-stud spaces. For both “LOS” and “WALL” collection locations the log periodic antenna was aligned with the main beam pointing down an office environment hallway at the collection device locations shown in Fig. 3.2. The collected Signal to Noise Ratio (SNR) over the Region Of Interest (ROI) was found to be SNR_C≈50, 40, 30 dB for “CAGE”, “LOS”, and “WALL” locations, respectively.

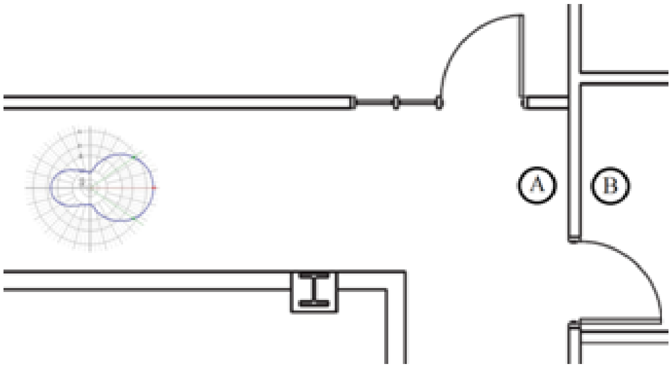


Figure 3.2: Operational indoor collection geometry showing collection receiver antenna pattern and ZigBee device (A) “LOS” and (B) “WALL” experimental collection locations.

3.2 Post-Collection Processing

The post-collection processing here was performed similarly to the methodology used in [12, 31]. The Agilent receiver collection files (*.cap* format) were converted for use with MATLAB® (*.mat* format) and post-collection processed by 1) detecting individual bursts using an amplitude-based threshold detection process, 2) removing detected bursts from the collection file, 3) down-converting individual bursts and applying baseband digital filtering, and 4) power scaling noise to achieve the desired SNR and model the effects of differing channel conditions. The Additive White Gaussian Noise (AWGN) was digitally filtered the same as collected bursts and power-scaled to achieve the desired SNR=[0-30] dB. Given the high collection SNR_C over the ROI, the like-filtered AWGN was added directly to the collected IQ data and was the dominant noise source.

3.2.1 Burst Detection.

The CC2420 devices were programmed to transmit bursts at a rate of approximately 14 bursts-per-second (1 burst every 69 ms) and transmissions were collected from one device at a time. The Aglient receiver stored the collected transmissions in a .cap file format which was converted to a .mat file for use in MATLAB®. Detection and extraction of burst responses were found using a amplitude-based threshold detection process with specific parameters including: termination threshold (t_T), detection threshold (t_D), minimum burst length (P_{MIN}), and maximum burst length (P_{MAX}). The instantaneous amplitude response ($a[n]$) of collected ZigBee bursts was calculated using (2.1) and converted to dB using,

$$a[n]_{dB} = 20\log_{10}\frac{a[n]}{1.0\ v}. \tag{3.1}$$

The result of (3.1) is illustrated in Fig. 3.3 for a collection containing N_B=4 bursts and a typical burst detection termination theshold (t_T). Burst detection begins by finding the global peak amplitude response C_G=max{|$a[n]$|} ∀ n in a given (.mat) collection file. Detection threshold t_D is then applied as shown in Fig. 3.4 to determine the leading and trailing edges of a declared burst, these edges correspond to leading/trailing edge sample indices (n_l, n_t) within $a[n]$ at which $|a[n]|{\approx}C_G - t_D$ occurs. The estimated burst duration ($n_t - n_l$) is calculated and compared to P_{MIN} and P_{MAX} to determine if the declared burst meets the estimated ZigBee pulse width, $P_{MIN}{<}(n_t - n_l){<}P_{MAX}$. If the declared burst meets all requirements, it becomes a *detected* burst and is extracted (removed from the collection file); else, the declared burst is discarded. This iterative process continues by finding the next maximum peak amplitude value $C_{MAX}{=}\max\{|a[n]|\}$, estimating burst duration, and so on. The detection process is terminated when either 1) the desired number of bursts are detected, or 2) the condition $C_{MAX}{<}C_G{-}t_T$ occurs for a declared burst indicating $\max\{|a[n]|\}$ is below the pre-established termination threshold, t_T. The specific values used for ZigBee burst detection are provided in Table 3.1.

Table 3.1: Burst detection parameters for ZigBee transmission collections.

Parameter	Variable	Value
Termination Threshold	t_T	6.0 dB
Detection Threshold	t_D	9.0 dB
Pulse Min Duration	P_{MIN}	850 μsec
Pulse Max Duration	P_{MAX}	870 μsec

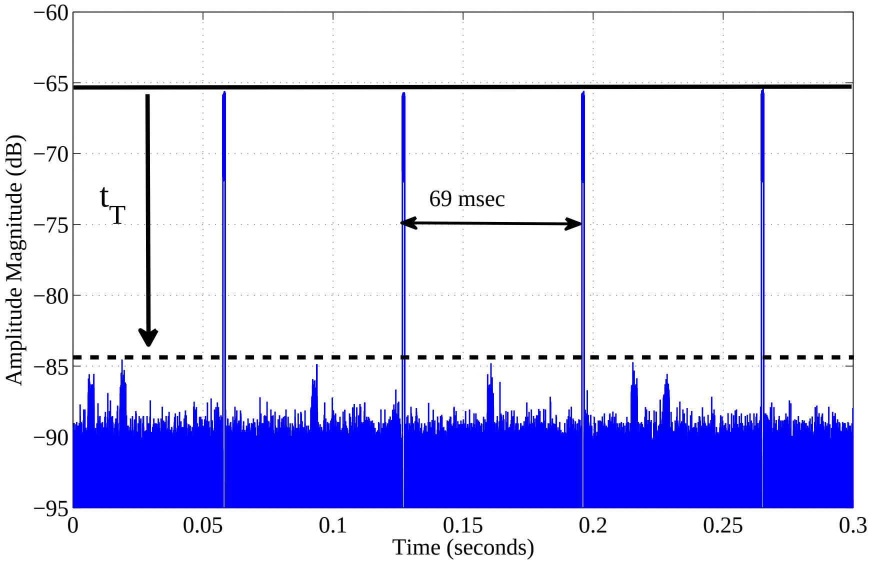


Figure 3.3: Representative ZigBee collection showing N_B=4 bursts and a typical processing *termination* threshold (t_T).

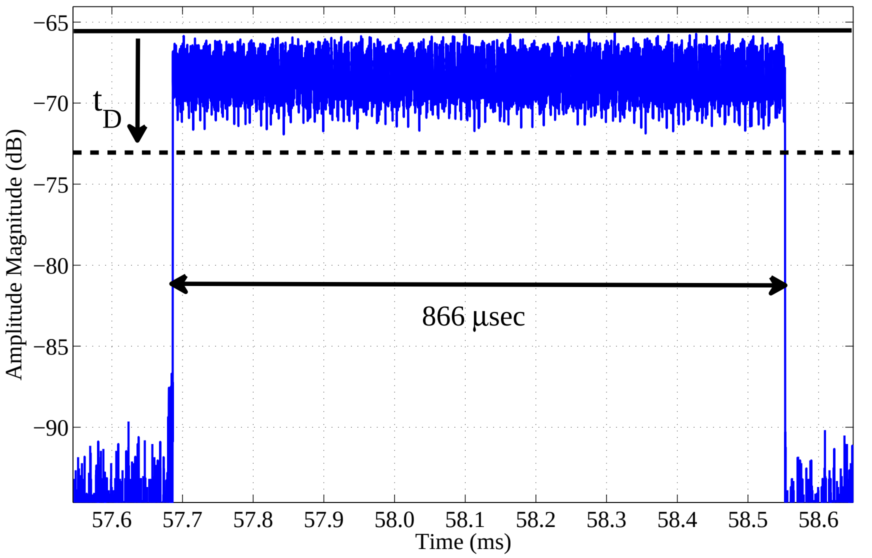


Figure 3.4: Representative *detected* ZigBee burst and typical *detection* threshold (t_D).

3.2.2 Digital Filtering.

The detected bursts are down-converted to baseband (f=0) using a Power Spectral Density (PSD) average estimated center frequency $\hat{f}_c$ for the 16 possible channels spanning 2.4 Ghz to 2.4835 GHz [26]. The down-conversion frequency ($\hat{f}_{DC}$) is estimated channel-by-channel such that bursts within estimated channels are all down-converted by the same estimated channel frequency. The down-converted signal is then digitally filtered using a 8th-order Butterworth baseband filter having a −3 dB bandwidth of W_{BB}=1.0 MHz. Figure 3.5 shows the PSD of a ZigBee baseband emission overlaid with the impulse response of the Butterworth baseband filter.

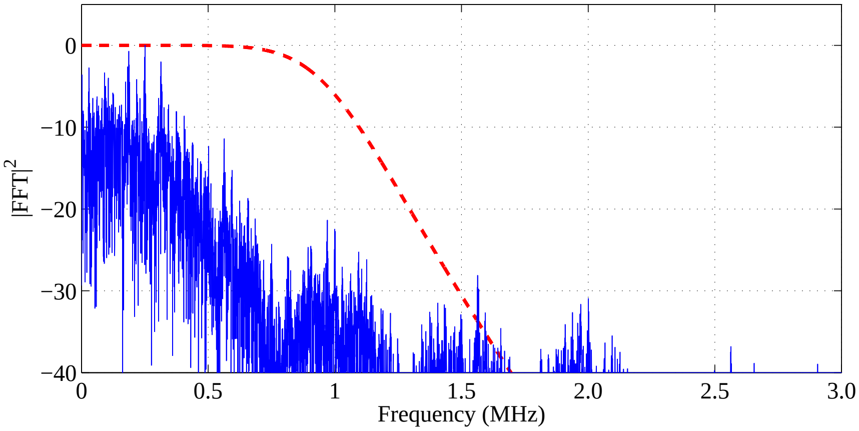


Figure 3.5: Representative ZigBee burst PSD response overlaid with an 8^{th}-order Butterworth digital filter impulse response.

3.2.3 Signal-to-Noise Ratio Scaling.

The high collected SNR_C over the ROI allows for the addition of power-scaled, like-filtered AWGN to generate analysis signals with SNR_A∈[0 30] dB. These analysis signals allow for classification and verification performance assessment under varying channel conditions. Using the analytic expression for an arbitrary complex sequence $\{x(i)\}$, i=1, 2, . . . , K, the estimated average power in X is given by,

$$X = \frac{1}{K}\sum_{i=1}^{K} x(i)x^{*}(i) \,, \tag{3.2}$$

where $x^{*}(i)$ is the complex conjugate of $x(i)$. The collected ZigBee signals are complex and consist of two components,

$$s_c(i) = s_t(i) + n_b(i) \,, \tag{3.3}$$

where $s_t(i)$ and $n_b(i)$ are the collected transmitted signal and collected background noise, respectively. The total power in s_c can be calculated as,

$$S_c = S_t + N_b \,, \tag{3.4}$$

where S_c was measured over the ROI and N_b was measured when no signal was present using (3.2) given by,

$$S_c = \frac{1}{K}\sum_{i=1}^{K} s_c(i)s_c^{*}(i) \,, \tag{3.5}$$

$$N_b = \frac{1}{K}\sum_{i=1}^{K} n_b(i)n_b^{*}(i) \,, \tag{3.6}$$

Rearranging (3.4) the transmitted signal power S_t is calculated and the estimated collected SNR in dB over the ROI is given by,

$$SNR_C^{dB} = 10 \times \log_{10}\left(\frac{S_t}{N_b}\right), \tag{3.7}$$

which yielded $SNR_C{\approx}50$, 40, 30 dB over the ROI region for “CAGE”, “LOS”, and “WALL” locations collections, respectively.

The desired scaled analysis signal $s_A(i)$ is generated by adding zero-mean, like-filtered, independent AWGN samples according to,

$$s_A(i) = s_t(i) + n_b(i) + n_G(i) \,, \tag{3.8}$$

where the average power in $\{n_G(i)\}$ is scaled to achieve a desired range of SNR_A.

A complex, zero-mean, normally distributed random sequence with an estimated average power of 1 (N_G=1) produces the AWGN samples. This complex sequence was digitally filtered by the same Butterworth filter used for the collected signal to produce like-filtered AWGN samples. The sequence is then power-scaled by R_n to achieve the desired SNR_A, with R_n calculated using,

$$R_n = \sqrt{10^{\frac{-SNR_A}{10}} \times S_t} \,, \tag{3.9}$$

which results in a total average AWGN power N_G given by,

$$N_G = \frac{1}{K} \sum_{i=1}^{K} R_n n_{AWGN}(i) R_n n^*_{AWGN}(i) \,. \tag{3.10}$$

The corresponding analysis SNR_A is then,

$$SNR_A^{dB} = 10 \times \log_{10}\left(\frac{S_t}{N_b + N_G}\right). \tag{3.11}$$

For general collection conditions the scaled AWGN power is generally much greater than the collected background noise power (N_G>>N_b) and (3.11) reduces to,

$$SNR_A^{dB} \approx 10 \times \log_{10}\left(\frac{S_t}{N_G}\right). \tag{3.12}$$

3.3 RF Fingerprint Generation

This section provides details on statistical time-domain RF-DNA fingerprint generation as introduced in Section 2.2. For this research, the ZigBee SHR region was selected as the ROI given that it 1) was experimentally observed within all bursts collected from all devices and 2) is independent of MAC frame information and payload data. The SHR region (40 bits total length) is comprised of a preamble sequence (32 bits in length) and the Start-of-Frame Delimiter (SFD) (8 bits in length) and consisted of 1920 collected time samples.

For this research the SHR time-domain signals were broken down into N_R=80 subregions (2 subregions for each bit) where 24 time samples were contained in each subregion. N_R=80 subregions was chosen because it showed improved device discrimination performance when compared to N_R=40 subregions (1 subregion for each bit). Full-dimensional RF-DNA fingerprints were generated using (2.7) through (2.14) based on N_C=3 signal characteristics (a, ϕ, f) and N_M=3 statistic metrics (σ^2, γ, κ), for a total of N_{Full}=N_M×(N_R + 1)×N_C=729 features per RF-DNA fingerprint. For this research the standard deviation statistic metric was omitted due to its close relation to variance. Figure 3.6 shows a representative time domain response for a ZigBee SHR region. The experimentally observed SHR duration of T_{SHR}≈160 μs is consistent with the IEEE 802.15.4 specification [26].

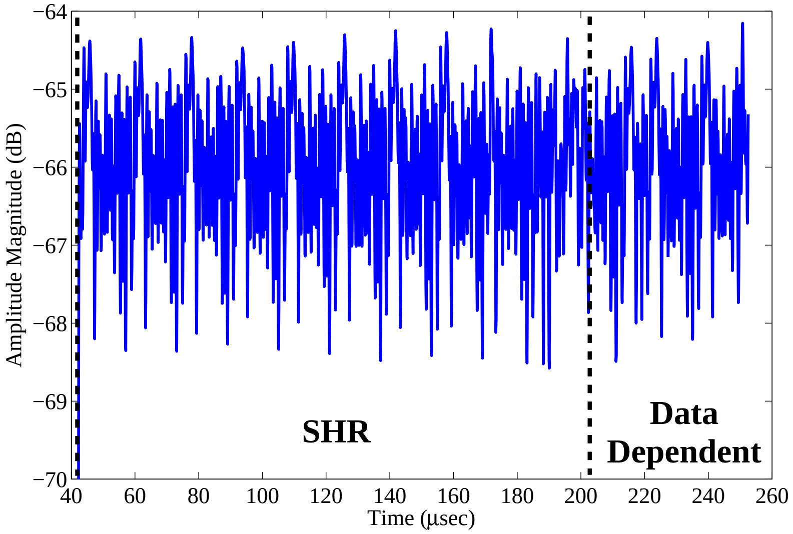


Figure 3.6: Representative ZigBee SHR response used as the region of interest for RF-DNA fingerprint generation.

3.4 Dimensional Reduction Analysis (DRA)

The Fisher-based MDA process in Section 2.3 inherently masks feature contribution to resultant classification performance and it is impossible to determine which features have the greatest impact. The goal of Dimensional Reduction Analysis (DRA) is to minimize the number of RF fingerprint features (N_F) while achieving a certain classification accuracy. One approach to minimize the number of features (dimensions) is to use the features that provide the most significant contribution to classification while removing less relevant features. Insight into feature relevance is addressed here quantitatively using: 1) a pre-classification KS-Test goodness-of-fit test [12, 31], and 2) a post-classification feature relevance ranking provided by GRLVQI processing [33, 36].

The KS-Test goodness-of-fit selection process includes [12, 31]:

1.) Generating a full-dimensional (N_F) feature set using (2.14) for N_{SHR} responses at a specific SNR from each of the N_D devices to be classified.

2.) Conducting N_{PW}=$[(N_D - 1)N_D]/2$ pairwise two-sample KS-tests using the N_F dimensional feature sets between every two devices under test, and forming a matrix of resultant p-values with dimension $N_{PW}{\times}N_F$.

3.) Summing each feature's p-values across pairwise combinations and rank-ordering the summed p-values from lowest-to-highest while tracking feature index number.

4.) Determining a summed p-value cutoff threshold, or arbitrarily setting a most relevant feature length l, to decide which features are retained for classification.

The quantitative pre-classification feature reduction process can be used to identify and select a most relevant, length l, subset of the full-dimensional RF-DNA feature set $\mathbf{F}$ prior to Multiple Discriminate Analysis, Maximum Likelihood (MDA/ML) classification. The KS-Test is a suitable option for analyzing statistical features differences and is used here to quantify differences in Cumulative Distribution Functions (CDF) between full-dimensional RF-DNA features from two devices. KS-Test results in Section 4.3

are presented as summed p-values from all pairwise combinations of the N_D devices considered, where lower p-values indicate a more significant data set difference [31].

The second alternative to feature selection is based on GRLVQI processing which inherently provides an indication of feature relevance following model development. The process here was adopted entirely from previous demonstrations showing that GRLVQI is a powerful tool for performing device classification and DRA [33, 38]. The GRLVQI process provides a relevance indicator (λ_i value) for each feature comprising the RF-DNA fingerprint at a specified SNR. The relevance value provides a measure of contribution to class (device) separation within the GRLVQI classification process. The higher the relevance value, the greater the impact on class separation. Feature DRA is achieved rank-ordering λ_i values and selecting the top-ranked, arbitrary length l, features from the full-dimensional feature set.

3.5 Device Discrimination Process

Statistical RF-DNA fingerprints for ZigBee device SHR responses are used as inputs into a device discrimination process. Figure 3.7 shows a block diagram for the device discrimination process used in this research. This process begins with separating collected RF fingerprints into *"Training"* and *"Testing"* sets, where the *"Training"* fingerprints are used for Multiple Discriminant Analysis (MDA) model development. Once a model is developed, *"Testing"* fingerprints are projected into the mapped feature space and used for either 1) *Device Classification* (a 1 vs. N_D "Looks most like?" assessment) or 2) *Device ID Verification* (a 1 vs. 1 "Looks how much like?" assessment).

3.5.1 MDA Model Development.

As introduced in Section 2.3, MDA can be applied when discrimination of two or more classes (devices) is required (N_D>2). For results presented in Chapter 4, MDA model development is performed using a pool of RF-DNA fingerprints from N_D=4 ZigBee devices

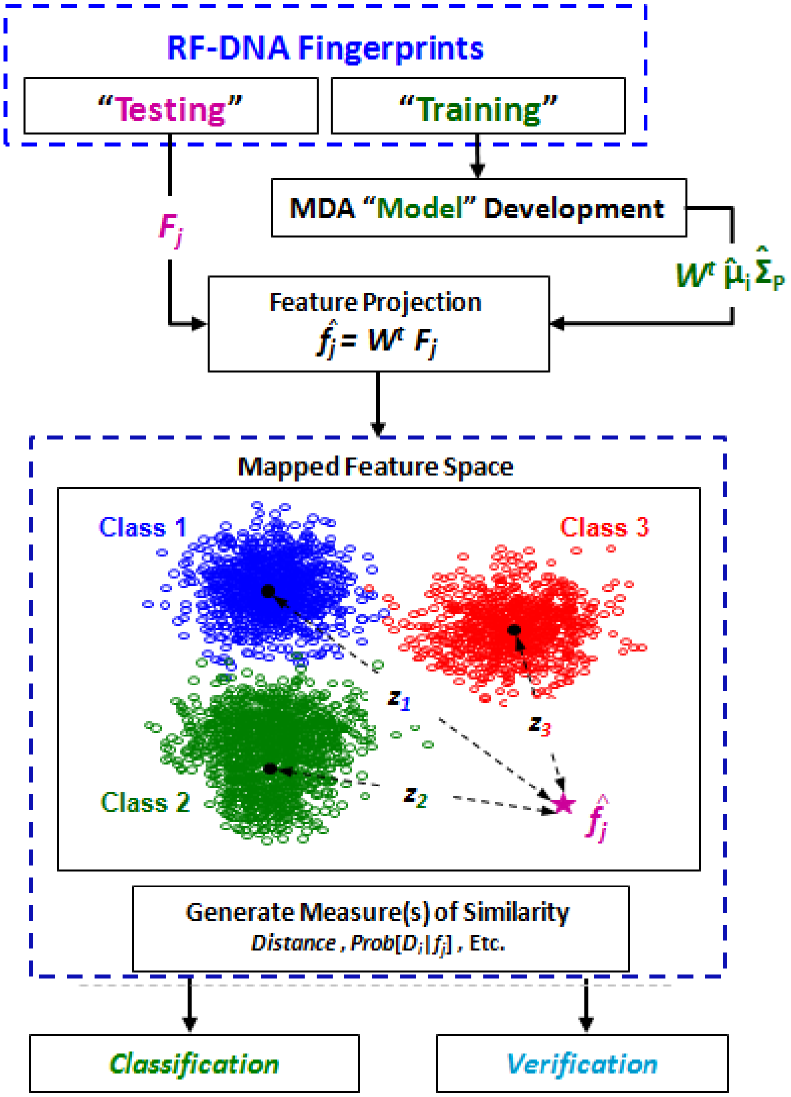


Figure 3.7: Block diagram of device discrimination process supporting both classification and verification using selected measures of similarity and test statistics.

(Dev1, Dev2, Dev3, and Dev4) constructed as a "hybrid" data set of fingerprints from the "CAGE", "LOS", and "WALL" collection scenarios; the result is referred to as a "hybrid" MDA model throughout the document. During model development MDA reduces input feature dimensionality by projecting N_F fingerprint features onto a (N_D-1)-dimensional subspace. The MDA projection matrix $\mathbf{W}^t$ is developed as shown in Fig. 3.8 using an iterative K-fold training process with a goal toward projecting higher-dimensional input fingerprint $\mathbf{F}$ data into a lower dimensional subspace such that inter-class separation is maximized and intra-class spread is minimized [13]. The parenthetical SNR denotes that the $\mathbf{W}^t(SNR)$, $\hat{\mu}_i(SNR)$, and $\hat{\Sigma}_P(SNR)$ generally varies with SNR, requiring MDA models to be developed for each SNR.

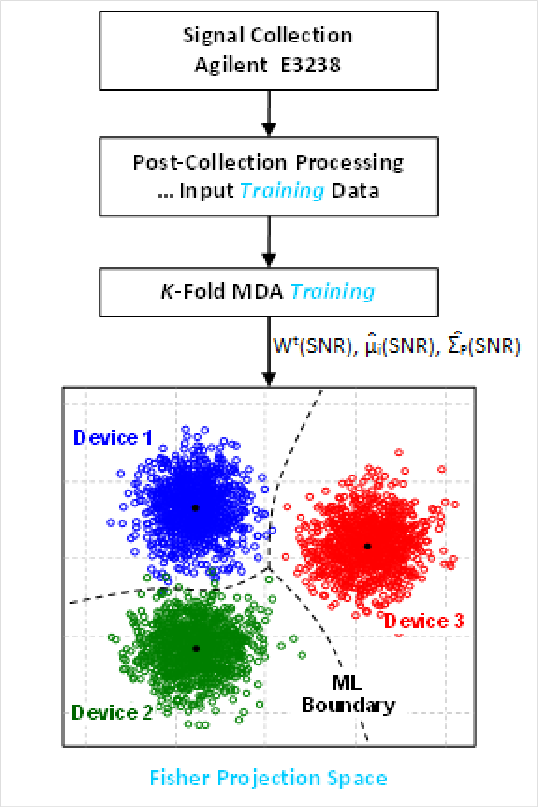


Figure 3.8: Signal collection, post-collection and K-fold MDA model development (training) processes. A representative 2D Fisher space is shown for N_D=3 ZigBee devices operating at SNR=10 dB. Clustering of the 100 projected training fingerprints (o) per device shown relative to class means (•).

For all results presented in Chapter 4, MDA model development was accomplished by using a K-fold cross-validation training process, shown in Fig. 3.9, where values of K=5 and K=10 are commonly used and provide sufficient statistical certainty [25]; a value of K=5 was used here. The K-fold training process consists of:

1. Randomly Parsing *"Training"* fingerprints into K blocks.
2. Separating K blocks such that K-1 blocks are used for training and one block is retained for model validation.
3. Performing MDA transformation on K-1 blocks using projection matrix $\mathbf{W}_K$, as described in Section 2.3.
4. Computing training class (device) means ($\hat{\mu}_i$) and pooled covariances ($\hat{\Sigma}_P$) to be used for Multivariate Gaussian (MVG) distributed models, as described in Section 2.4.
5. Tracking fold ML classification performance ($\%C_K$) using the retained validation block, as described in Section 2.4.
6. Repeating steps 2-5 such that a different block is retained for validation until K iterations are completed.
7. Determining the $\mathbf{W}_K$ and corresponding $\hat{\mu}_i$, and $\hat{\Sigma}_P$ that achieved maximum (Best) classification performance (highest $\%C_K$).

3.5.2 Device Classification.

Once MDA model development is accomplished, device classification is performed using a Maximum Likelihood (ML) classifier as described in Section 2.4, with input *"Testing"* fingerprints classified as being affiliated with one of N_D=4 possible devices. For ML classification, the prior probabilities are assumed to be equal, the costs uniform, and the device likelihoods have a MVG distribution with means ($\hat{\mu}_i$) and covariances ($\hat{\Sigma}_P$) as computed during MDA model development. The ML classification process consists of: 1) inputting a *"Testing"* fingerprint $\mathbf{F}_j$ for a collected emission from an unknown device

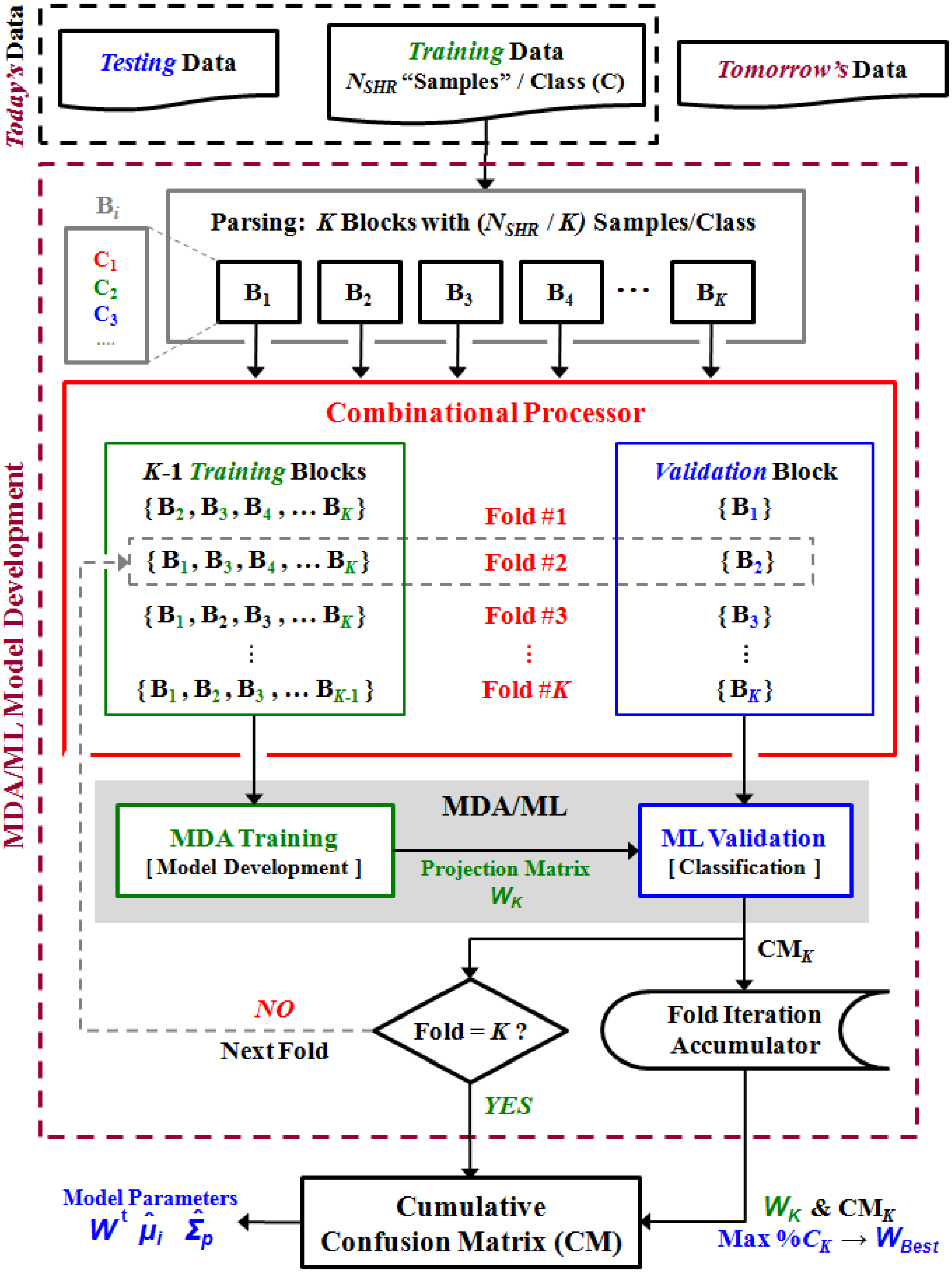


Figure 3.9: Illustration of K-fold cross-validation training process used for MDA model development. The "best" model $\mathbf{W}_B$ is selected as the $\mathbf{W}_K$ yielding maximum $\%C_K$.

D_j, 2) projecting $\mathbf{F}_j$ into the Fisher space using $\hat{\mathbf{f}}_j = \mathbf{W}^t\mathbf{F}_j$, and 3) associating $\hat{\mathbf{f}}_j$ as being from the device with the maximum conditional likelihood probability according to,

$$D_i \ : \ \arg\max_i \Big[\ p(D_i|\hat{\mathbf{f}}_j) \ \Big] \tag{3.13}$$

where i=$1, 2, \ldots, N_D$ and $p(D_i|\hat{\mathbf{f}}_j)$ is the conditional likelihood probability that fingerprint $\hat{\mathbf{f}}_j$ belongs to device D_i. Correct classification is achieved when projected *"Testing"* fingerprints are classified to be from their true device. Average percent correct (%C) device classification is calculated as the percentage of the time the classifier correctly assigns the fingerprint to its true device over all trials.

3.5.3 Device ID Verification.

For device ID verification (a 1 vs. 1, claimed vs. actual, "Looks how much like?" assessment), the process used here is consistent with the methodology used in [11, 12, 33]. The focus here is on answering "Does the device's current RF-DNA fingerprint match the stored RF fingerprint template associated with its claimed bit-level identity?". RF-DNA fingerprints can be used to authenticate a device's claimed bit-level identity, i.e., a device wants to access a network and has presented its MAC address, SIM number, IMEI number, etc., to gain access [11]. Bit-level credentials can be easily replicated by rogue devices, and RF-DNA fingerprint verification provides a means to mitigate unauthorized access attempts. This is done by a 1-to-1 comparison of current vs. claimed RF signatures, with the claimed signature being a stored template associated with the claimed bit-level identity. Each designated authorized device in a network will have a stored RF signature reference template that is used when a current *"Testing"* RF fingerprint is received and has claimed an ID of a authorized device. The device ID verification process is used here for two performancwe assessments, including:

1. *Authorized Device ID Verification*: Granting network access to authorized devices presenting *true* bit-level credentials.

2. *Rogue Device Rejection*: Denying network access to unauthorized rogue devices presenting *false* bit-level credentials.

3.5.3.1 Authorized Device ID Verification.

Authorized device ID verification is an assessment of how similar a device's current RF fingerprint matches the stored reference model associated with the claimed identity, when only considering *"Testing"* RF fingerprints from a pool of N_D authorized devices. The similarity measure, or verification test statistic (z_V) reflects "How well" the current and claimed RF fingerprint identities match and is compared with a threshold (t_V) to verify the device's claimed ID and grant or deny network access. Verification test statistics (z_V) can be generated from probability-based measures or geometric measures such as distance, spatial angle, etc. The specific test statistics used here for *Device ID Verification* are inherently provided in the "posterior" output variable of MATLAB® `classify` function. The posterior matrix contains normalized conditional Multivariate Gaussian posterior probabilities given by,

$$z_V = \frac{p(D_i|\hat{\mathbf{f}}_j)}{\sum\limits_{k=1}^{N_D} p(D_k|\hat{\mathbf{f}}_j)} , \tag{3.14}$$

where $i = 1, 2, \ldots, N_D$ and $\hat{\mathbf{f}}_j$ is the current projected RF fingerprint claiming to have an ID from device, D_i. For this research it is assumed that each authorized device claims N_D IDs (one for each authorized device). For a given *"Testing"* RF fingerprint this produces N_D test statistics, where one test statistic is from the proper *true* device and $N_D - 1$ test statistics are from device's claiming a *false* ID.

Authorized device ID verification is evaluated one claimed ID at a time, where test statistics are generated for all N_D authorized device's *"Testing"* data set producing two Probability Mass Functions (PMFs): 1) an In-Class PMF, and 2) an Out-of-Class PMF. Where the In-Class PMF is formed by test statistics (z_V) from a device that is *actually* who it claims to be, the current RF fingerprint is from the proper authorized device. Each authorized device will have a corresponding In-Class PMF and these are known as the stored true reference templates associated with the authorized device's ID. Out-of-Class PMF is generated using (z_V) for the case when a authorized device *falsely* claims an identity of a different authorized device. Figure 3.10 shows a representative In-Class and Out-of-Class PMF generated from arbitrary test statistics (z_V) for a single claimed ID. The In-Class probability is defined as $p[z_V|C_i, D_j]$, where i=j and C_i is the claimed Device ID (i=$1, 2, \ldots, N_D$) and D_j is the actual (current) device. The corresponding Out-of-Class probability is denoted as $p[z_V|C_i, D_j]$, where i≠j and j=$1, 2, \ldots, N_D$.

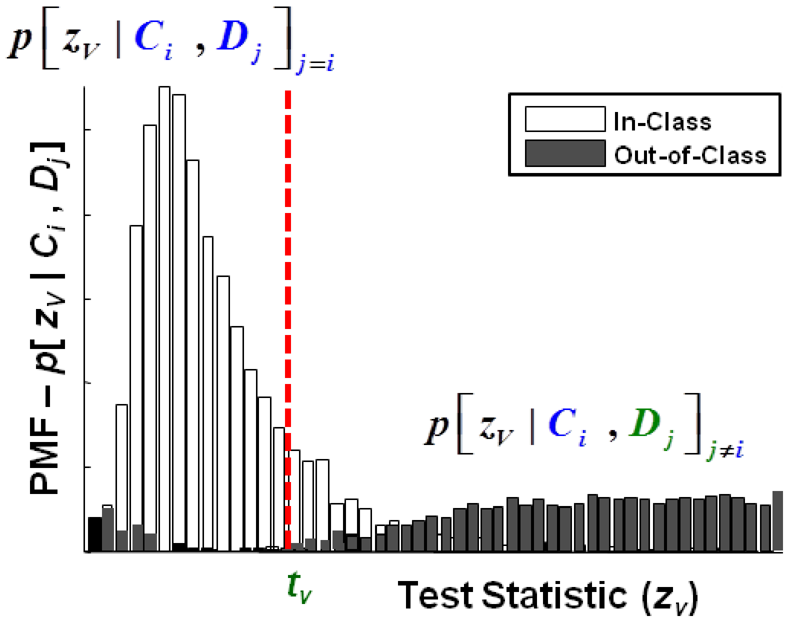


Figure 3.10: Representative In-Class (unfilled) and Out-of-Class (filled) Probability Mass Functions (PMFs) for an arbitrary test statistic (z_V). These are used to generate an *Authorized Device ID Verification* ROC curve for a specific claimed ID and varying threshold t_v.

Authorized device ID verification is evaluated for all claimed IDs and is assessed using conventional Receiver Operating Characteristics (ROC) curve analysis [15]. True and false device ID verification rates are generated by varying the threshold (t_V) shown in Fig. 3.10

and measuring the area of each PMF. True Verification Rate (TVR) is a measure of "how well" current RF fingerprints match its true claimed ID and is the area under the In-Class PMF when $z_V < t_V$. The corresponding False Verification Rate (FVR) provides a measure of "how well" current RF fingerprints match a false claimed ID and is the area under the Out-of-Class PMF when $z_V < t_V$. As the threshold (t_V) varies, corresponding TVR and FVR are used to generate a ROC performance curve. As shown in Fig. 3.11, ROC performance is a function of SNR. Representative thresholds ($t_1 < t_2 < t_3$) are shown to emphasis that a given verification threshold t_V dictates TVR and FVR performance.

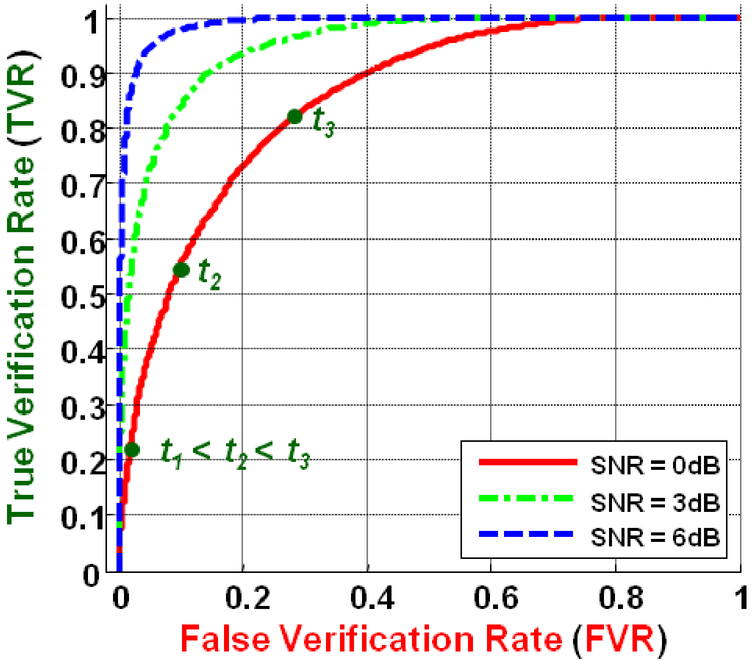


Figure 3.11: Representative *Authorized Device ID Verification* ROC curves showing performance variation as a function of SNR, i.e., degradation for decreasing SNR.

3.5.3.2 Rogue Device Rejection.

Using the same process as authorized device ID verification, *Rogue Device Rejection* capability can be measured when a rogue device presents false bit-level credentials in an attempt to gain unauthorized network access. Rogue device rejection is an assessment of how similar unauthorized rogue device's current RF fingerprint matches the stored true reference template associated with the claimed identity presented by the rogue device. *"Testing"* RF fingerprints are generated for previously unseen N_R rogue devices using the same method describe in this chapter and projected into the (N_D-1) Fisher subspace. The z_V test statistics from (3.14) are generated to provide a measure of "How well" the rogue device's current RF fingerprint matches claimed authorized devices RF fingerprint. For this research it is assumed that each rogue device claims N_D IDs (one for each authorized device). For a given rogue *"Testing"* RF fingerprint this produces N_D test statistics, where the rogue device claimed a *false* ID.

Rogue device rejection is evaluated one claimed ID at a time, where test statistics are generated for a single N_R rogue device's *"Testing"* data set producing a new Out-of-Class PMF, that is compared to the stored true reference template (In-Class PMF) associated with the rogue device's claimed ID. For a single claimed ID, Fig. 3.12 shows a representative *unchanged* In-Class PMF from Fig. 3.10 and the new Out-of-Class PMF generated from arbitrary test statistics (z_V). The In-Class probability is defined as $p[z_V|C_i, D_j]$, where i=j and C_i is the claimed Device ID (i=$1, 2, \ldots, N_D$) and D_j is the actual (current) device. The corresponding Out-of-Class probability is denoted as $p[z_V|C_i, D_k]$, where k≠j and k≠$1, 2, \ldots, N_D$, and D_k is a rogue device.

Rogue device rejection is assessed using conventional ROC curve analysis [15]. Varying the threshold (t_V) shown in Fig. 3.12 and measuring the area under the curve for each PMF will determine the True Verification Rate (TVR) and Rogue Accept Rate (RAR). TVR is a measure of "how well" current RF fingerprints match its true claimed ID and is the area under the In-Class PMF when z_V<t_V. The area under the In-Class-PMF is the same as shown in Fig. 3.10. The corresponding RAR provides a measure of "how well" current rogue RF fingerprints match a falsely claimed authorized device ID and is the area under

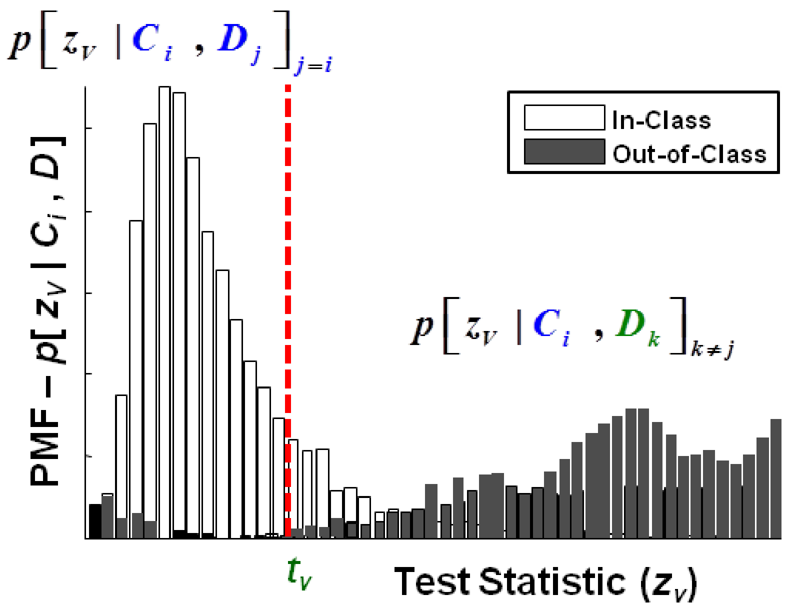


Figure 3.12: Representative In-Class (unfilled) PMF from Fig. 3.10 and Out-of-Class (filled) PMF for arbitrary test statistic z_V. These are used to generate an *Rogue Device Rejection* ROC curve for a specific claimed ID and selected threshold t_v.

the Out-of-Class PMF when $z_V < t_V$. The RAR is a measure of "how often" a rogue device is granted network access when falsely claiming a bit-level identity of a authorized network device. Rogue Reject Rate (RRR) is defined as $RRR = 1 - RAR$; a higher RAR (lower RRR) reflects poorer security performance. Figure 3.13 shows representative authorized device ID verification and rogue device rejection ROC performance curves, illustrating the process of setting a threshold (t_V) to achieve a desired TVR corresponds to a given authorized device false verification rate and a rogue accept rate for a specific claimed ID.

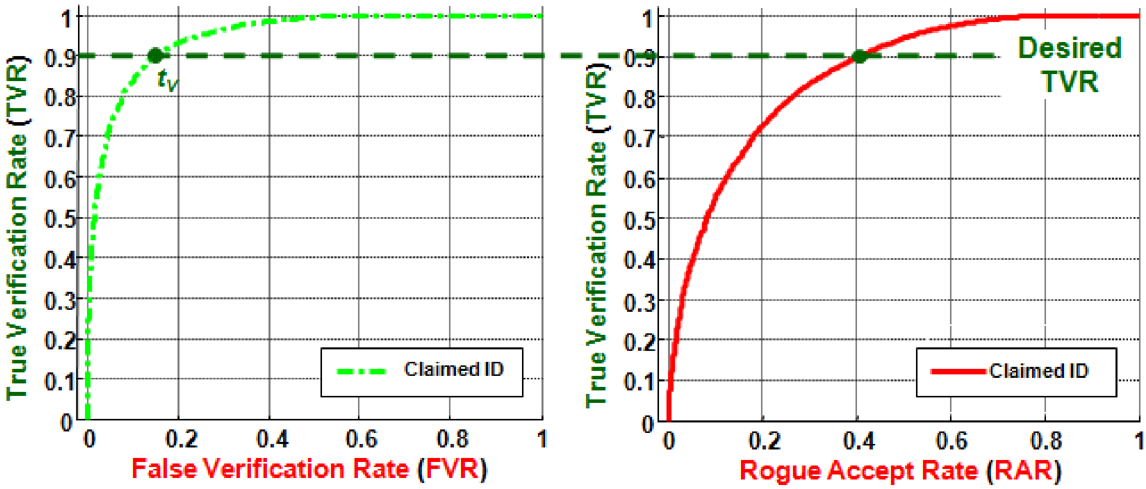


Figure 3.13: Representative *Authorized Device ID Verification* and corresponding *Rogue Device Rejection* ROC curves. Verification threshold t_V is set to achieve desired authorized device TVR and FVR which maps directly to a corresponding rogue device RAR (RRR) for a specific claimed ID.

IV. Results and Analysis

This chapter provides results for ZigBee device discrimination, to include *Device Classification* and *Device ID Verification* using full-dimensional and reduced dimensional RF-DNA feature sets. The reduced dimensional subsets are obtained through Dimensional Reduction Analysis (DRA) as described in Sect. 3.4 using a *qualitative* phase-only feature selection process as in [11, 31] and two *quantitative* selection methods, including: 1) pre-classification Kolmogorov-Smirnov (KS)-Test *p*-value ranking and 2) post-classification Generalized Relevance Learning Vector Quantization-Improved (GRLVQI) feature relevance ranking. *Device Classification* and *Device ID Verification* are performed using the methodology discussed in Section 3.5. Section 4.1 provides the details how Multiple Discriminant Analysis (MDA) training was accomplished, including the selection of *Training* and *Testing* data sets. Section 4.2 provides baseline Multiple Discriminate Analysis, Maximum Likelihood (MDA/ML) classification performance using full-dimensional RF-DNA fingerprints. Section 4.3 provides comparative DRA feature selection results for the three selection methods considered. Section 4.4 provides *Device Classification* results using selected DRA feature sets, and Section 4.5 provides verification results, including *Authorized Device ID Verification* and *Rogue Device Rejection* performance using DRA reduced feature sets.

4.1 MDA Training and Model Development

MDA training was accomplished using N_{SHR}=500 independent ZigBee Synchronization Header (SHR) responses collected from each location (“CAGE”, “LOS”, and “WALL”) for each device used for hybrid model development (Dev1, Dev2, Dev3, Dev4). In addition, N_{Nz}=5 independent, like-filtered, Monte Carlo Noise realizations were added to the SHR responses for each analysis SNR considered. Thus, for N_D=4 devices MDA training, *K*-

fold generation of the "best" MDA model ($\mathbf{W}^t$, $\hat{\mu}_i$, $\hat{\Sigma}_P$) and MVG statistics of projected *Training* fingerprints, are on a total of N_{TNG}=(500 SHR)×(3 Locations)×(5 N_{Nz})=7500 independent *Training* realizations per device. Results for classification are likewise based on N_{SHR}=500 *Testing* fingerprints per location for each device and N_{Nz}=5 noise realizations per SNR, resulting in N_{TST}=7500 *Testing* realizations. This large number of trials reduced the CI=95% Confidence Interval (CI) bars to within the vertical extent of the plotted data markers. Therefore, the CI=95% are intentionally omitted in all plots to enhance visual clarity and qualitative assessment.

4.2 Device Classification: Full-Dimensional Performance

Full-Dimensional RF-DNA feature sets are based on N_C=3 signal characteristics (a, ϕ, and f), N_M=3 statistics (σ^2, γ, and κ), and $N_R + 1$=81 total regions. Thus, the composite fingerprint $\mathbf{F}$ for each collected emission is comprised of N_F=729 RF fingerprint features as given by (2.14). Figure 4.1 shows the full-dimensional classification *Testing* performance for the hybrid location (responses from "CAGE", "LOS", and "WALL") scenario and $SNR \in [0\ 24]$ dB. An arbitrary performance benchmark of $\%C$=90% (average across devices) is achieved at SNR=9.2 dB($\approx$10.0 dB), with all devices achieving $\%C$=80% or better classification at this point. Each device classification performance curve shown in Fig. 4.1 is an average performance across locations ("CAGE", "LOS", and "WALL").

4.3 Device Classification: DRA Feature Selection

Results in Fig. 4.1 show that the arbitrary $\%C$=90% benchmark can be achieved for all devices at various SNR using a full-dimensional N_F=729 feature set, with average cross-device $\%C$=90% achieved at $SNR\approx$10.0 dB. Feature down-selection was next performed using DRA to determine the minimum number of features required to maintain average cross-device $\%C$=90%. Feature relevance was determined using RF fingerprints extracted from emissions at SNR=10.0 dB (the SNR at which $\%C$=90% in Fig. 4.1). Quantitative

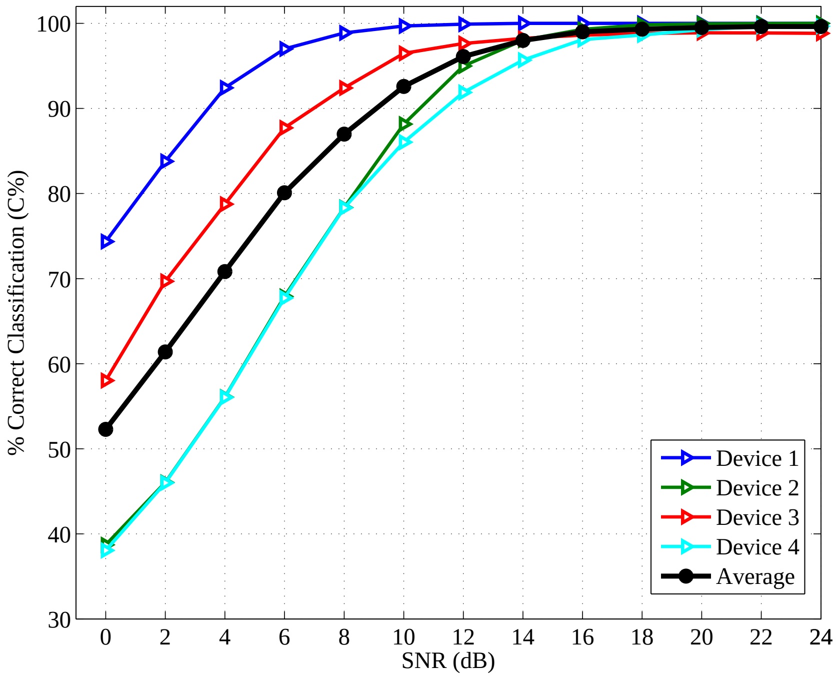


Figure 4.1: MDA/ML *Device Classification* performance using a ***full-dimensional*** (N_F=729) ZigBee feature set at indicated SNR. The cross-device average is shown and used for subsequent comparison with DRA performance results.

DRA was performed using the N_F=729 full-dimensional feature with 1) pre-classification KS-Test p-value ranking and 2) post-classification GRLVQI λ_i feature relevance ranking.

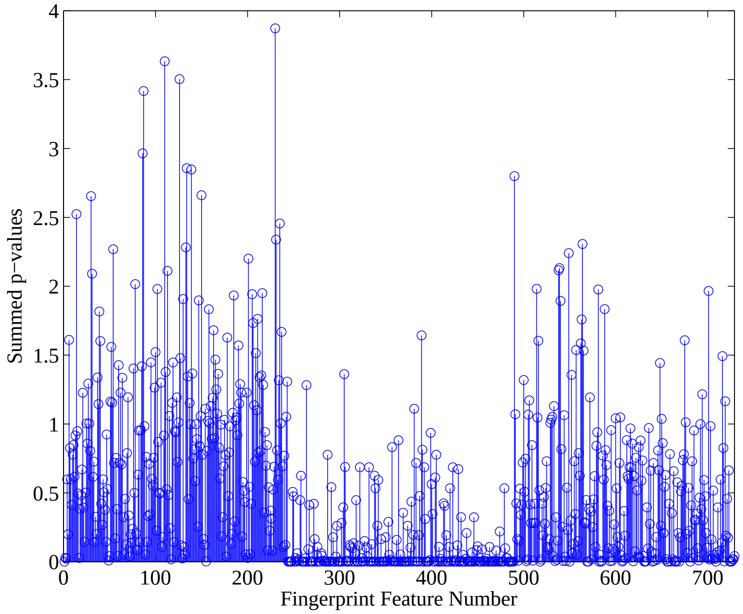
Quantitative DRA enables identification and selection of feature subsets, where the most relevant features are selected from the full-dimensional feature set. Figure 4.2 shows the N_F=729 full-dimensional ZigBee feature number indices and corresponding relevance indicators for SNR=10.0 dB using 1) pre-classification KS-Test p-values and 2) post-classification GRLVQI λ_i relevance values. Most significant feature relevance is indicated by a *lower* summed p-value from the KS-Test and a *higher* λ_i from the GRLVQI process. The DRA process simply involves sorting Fig. 4.2 results to establish a rank-ordering that can be used to select a desired number of most relevant features.

4.4 Device Classification: DRA Performance

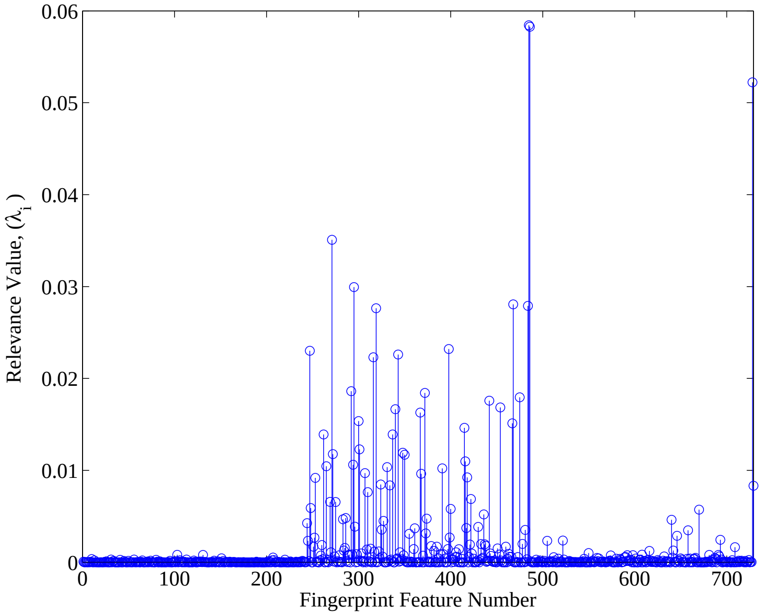
Previous research [11, 31] has qualitatively shown that ZigBee phase-derived features possess greater discriminating information than either amplitude-derived or frequency-derived features when used with an MDA/ML classifer. As detailed in Section 3.3, the full-dimensional ZigBee feature set consists of N_F=729 total features, including $N_{\widehat{F}}$=243 amplitude, phase, and frequency features. Figure 4.3 displays DRA subsets comprised of $N_{\widehat{F}}$=243 selected features and their corresponding indices for 1) qualitative phase-only feature selection, 2) quantitative KS-Test top-ranked feature selection, and 3) quantitative GRLVQI top-ranked feature selection.

Figure 4.4 shows average *Device Classification* performance using the N_F=729 full-dimensional feature set and the DRA≈66% subsets ($N_{\widehat{F}}$=243 features retained) shown in Fig. 4.3. Relative to full-dimensional performance, the DRA≈66% subsets yield relatively consistent classification performance and exhibit a "gain" of $G \approx -1.0$ dB at the $\%C$=90% benchmark; the "gain" metric is introduced here for comparative assessment and defined as the difference, expressed in dB, in required SNR (dB) for two systems, methods, etc., to achieve a specified performance $\%C$.

Further reduction of RF-DNA fingerprint dimensionality is obtained using the top-ranked $N_{\widehat{F}}$=200, 100, 50, and 25 features that were quantitatively selected using the 1) pre-classification KS-Test and 2) post-classification GRLVQI relevance rankings. Figure 4.5 displays the top-ranked $N_{\widehat{F}}$=243, 200, 100, 50, and 25 features from both quantitative DRA methods and their corresponding index number within the full-dimensional feature set.



(a) KS-Test: Lower → Greater Relevance



(b) GRLVQI: Higher → Greater Relevance

Figure 4.2: Unsorted DRA feature relevance indicators: (a) KS-Test p-values and (b) GRLVQI λ_i relevance values. Results shown here for SNR=10.0 dB which corresponds to a cross-device %C≈90% in Fig. 4.1.

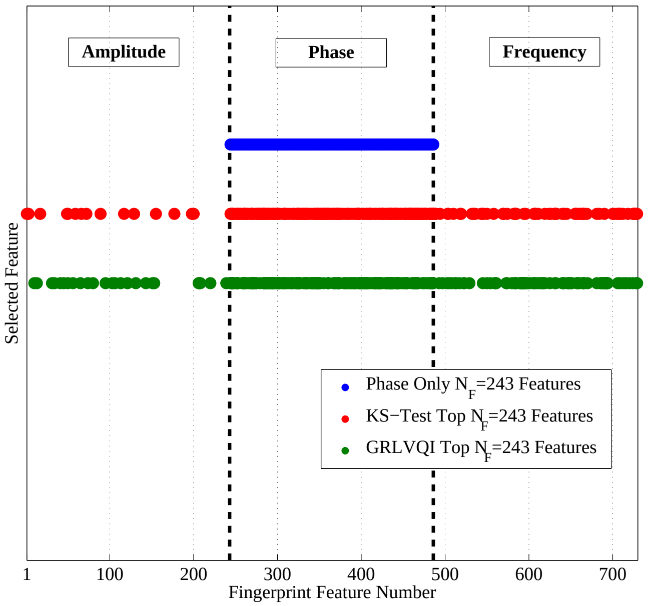


Figure 4.3: DRA Selected $N_{\widehat{F}}$=243 subsets of full-dimensional (N_F=729) feature set. Selection based on 1) qualitative phase-only, 2) quantitative top-ranked KS-Test, and 3) quantitative top-ranked GRLVQI feature selection methods.

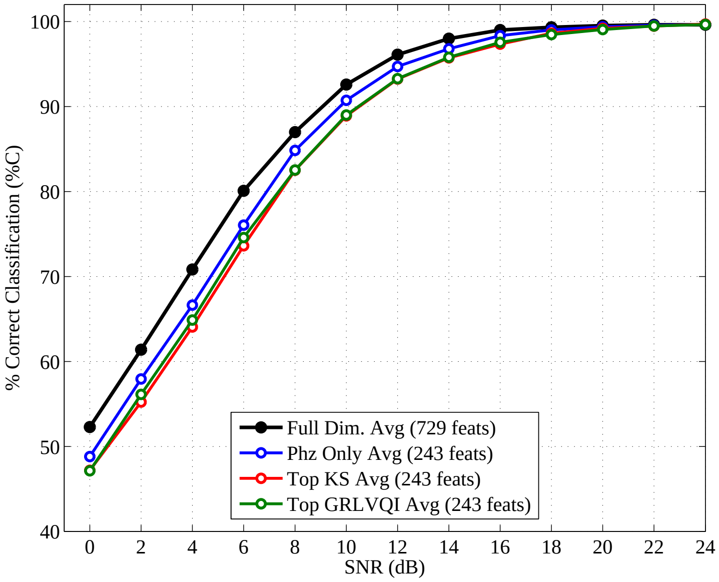
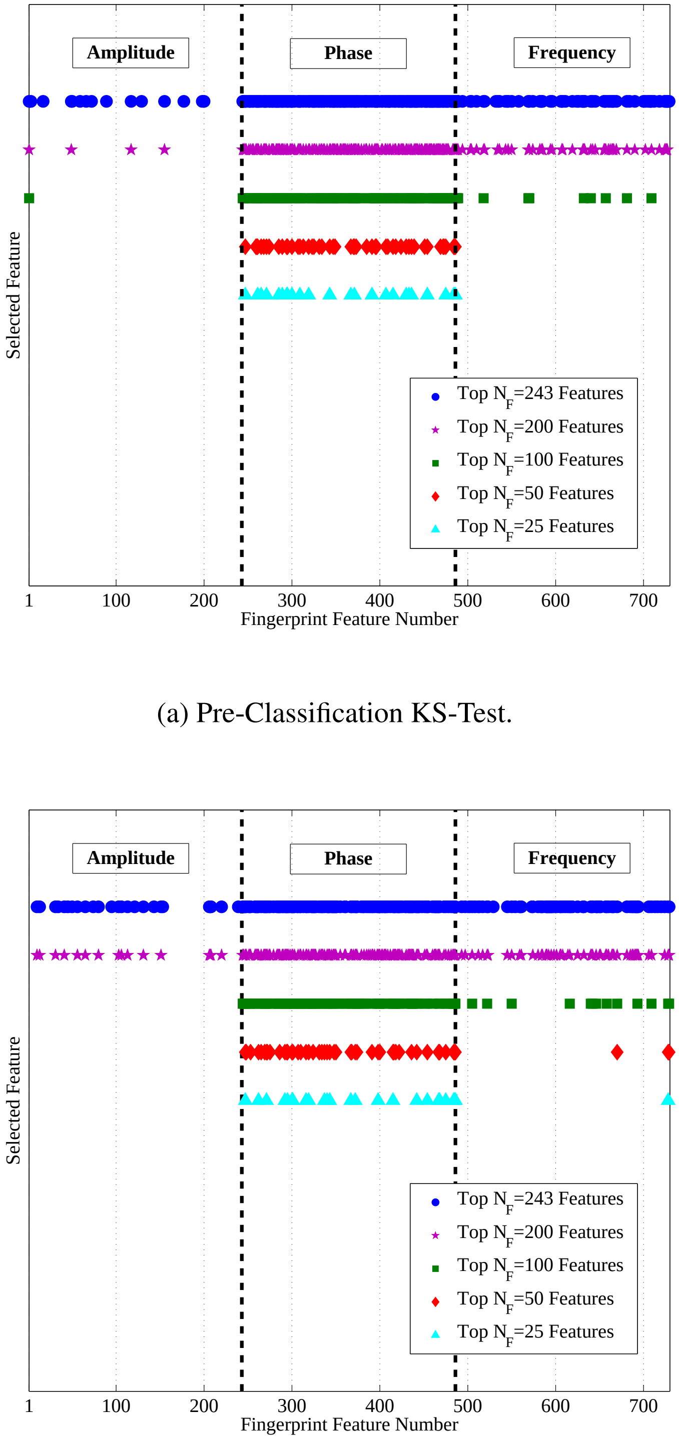


Figure 4.4: Average MDA/ML device classification performance using DRA selected $N_{\widehat{F}}$=243 feature subsets shown in Fig. 4.3. Full-dimensional N_F=729 performance from Fig. 4.1 provided for comparison.



(a) Pre-Classification KS-Test.

(b) Post-Classification GRLVQI.

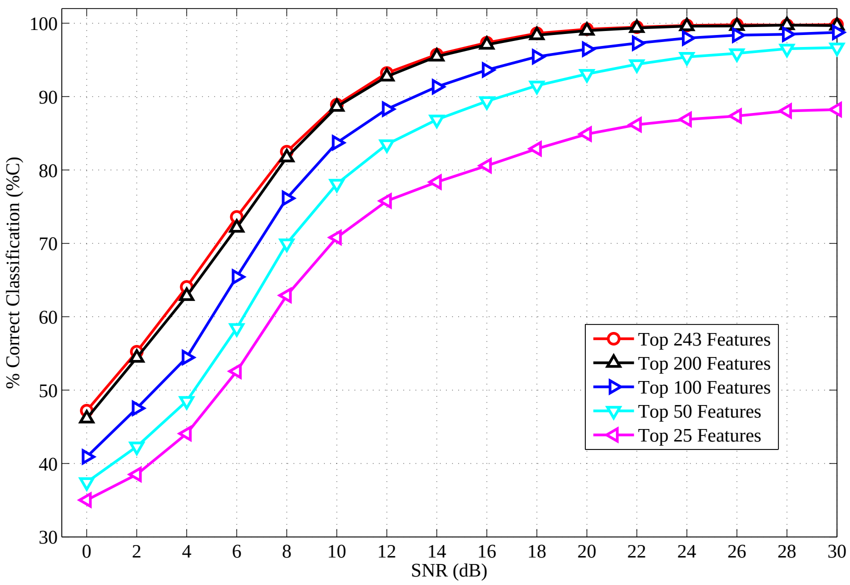
Figure 4.5: Illustration of top-ranked $N_{\widehat{F}}$=243, 200, 100, 50, and 25 DRA feature subsets using (a) pre-classification KS-Test and (b) post-classification GRLVQI rankings.

The effect of additional feature reduction and assessment of hybrid location classification performance is shown in Fig. 4.6 using DRA subsets containing the top-ranked $N_{\widehat{F}}$=243, 200, 100, 50 and 25 features that were quantitatively selected using 1) pre-classification KS-Test and 2) post-classification GRLVQI relevance rankings. Considering the previously established $\%C$=90% benchmark for assessing DRA classification performance, results in Fig. 4.6 show that:

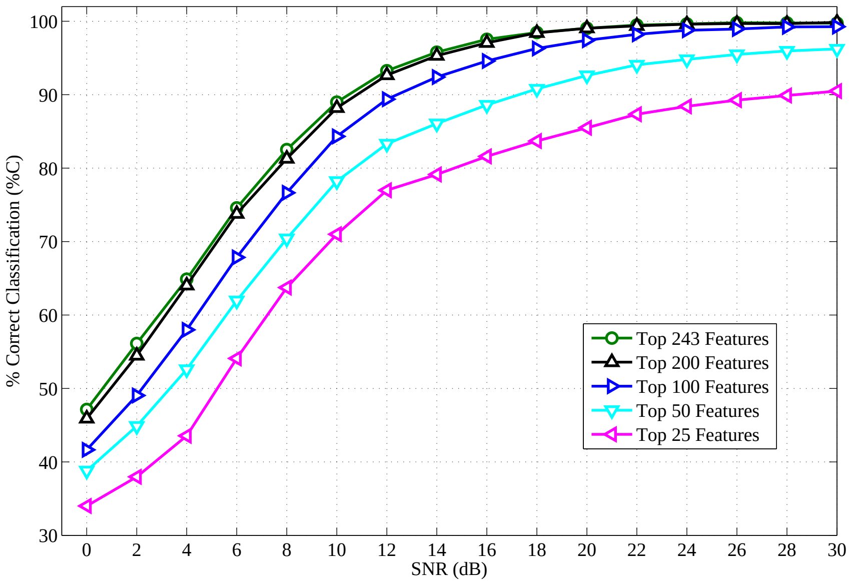
1. The required SNR for KS-Test top-ranked $N_{\widehat{F}}$=243 and $N_{\widehat{F}}$=50 feature sets approximately spans $SNR \in$[10 18] dB, with the top-ranked $N_{\widehat{F}}$=25 feature set never achieving the %C=90% benchmark. This is an indication that the MDA model development process is unable to achieve adequate inter-class separation and/or sufficient intra-class spread minimization using only $N_{\widehat{F}}$=25 features.

2. The required SNR for GRLVQI top-ranked $N_{\widehat{F}}$=243 and $N_{\widehat{F}}$=50 feature sets approximately spans $SNR \in$[10 18] dB which is consistent with KS-Test feature selection performance. However, the GRLVQI top-ranked $N_{\widehat{F}}$=25 feature set also achieves the %C=90% benchmark at $SNR \approx$30 dB.

The KS-Test and GRLVQI feature selection performances in Fig. 4.6 are summarized in Table 4.1 which shows the “Gain” for each DRA case relative to performance using the DRA≈66% reduced $N_{\widehat{F}}$=243 feature set.



(a) KS-Test Feature Selection.



(b) GRLVQI Feature Selection.

Figure 4.6: MDA/ML *Device Classification* performance using DRA subsets from Fig. 4.5 selected by (a) KS-Test p-values and (b) GRLVQI λ_i relevance values. Average $N_{\widehat{F}}$=243 DRA performance from Fig. 4.4 provided for comparison.

Table 4.1: MDA/ML *Device Classification* performance "Gain" (dB) for DRA subsets in Fig. 4.6 relative to performance using the DRA $N_{\widehat{F}}$=243 feature subset.

DRA Method	Number of DRA Features ($N_{\widehat{F}}$)				
	243	200	100	50	25
KS-Test	0.0 dB	-0.2 dB	-2.6 dB	-6.1 dB	N/A
GRLVQI	0.0 dB	-0.3 dB	-1.9 dB	-6.75 dB	-17.8 dB

4.5 Device ID Verification

Verification of a device's claimed bit-level ID provides a means for granting authorized devices network access while denying access to unauthorized devices. It is assumed here that a device wanting to gain network access provides a claimed bit-level ID and that RF-DNA features can be used to authenticate the claimed ID. The *Device ID Verification* process performs a 1-to-1 comparison between a device's *current* RF-DNA fingerprint and a *stored* reference fingerprint for the claimed bit-level ID. Device ID verification is accomplished here using the methodology described in Section 3.5.3 and emissions from 10 ZigBee devices, including: 1) the same N_D=4 authorized devices used previously for device classification assessment (Dev1, Dev2, Dev3, and Dev4), and 2) an additional N_R=6 unauthorized "rogue" devices (Dev5, Dev6, Dev7, Dev8, Dev9 and Dev10). The verification process is used to assess both *Authorized Device ID Verification* performance using the N_D=4 authorized devices, and *Rogue Device Rejection* performance using the N_R=6 rogue devices. Of particular importance is that the hybrid MDA model developed in Sect. 4.1 for *Device Classification* is also used here for verification assessment.

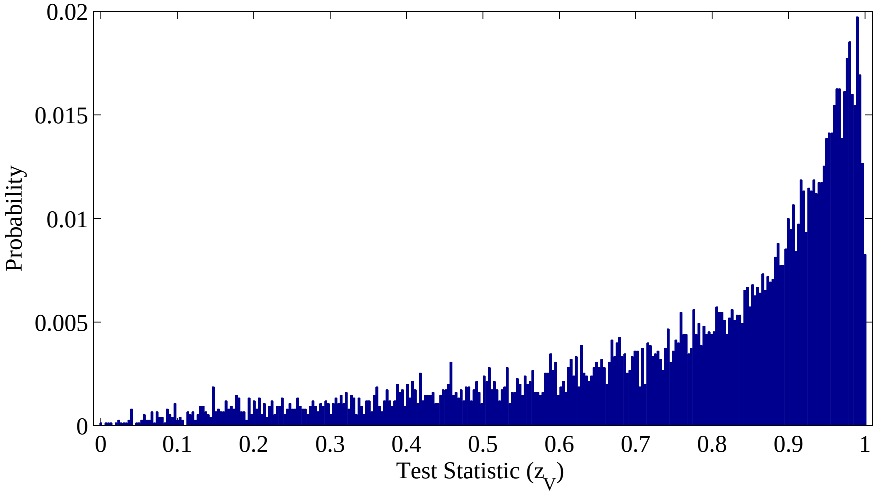
4.5.1 Authorized Device ID Verification.

Authorized device ID verification is performed using the same independent N_{TST}=7500 projected *Testing* fingerprints from classification for each of the N_D=4 authorized devices.

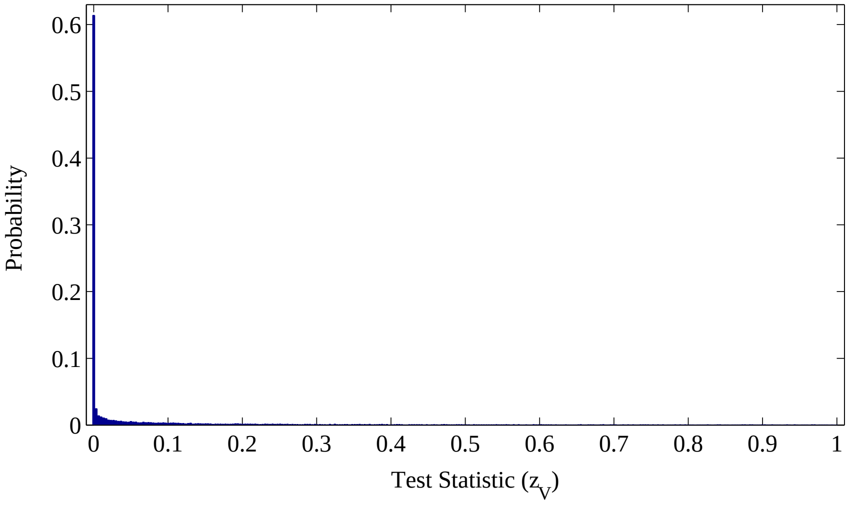
Verification performance is evaluated at SNR=18.0 dB using $N_{\widehat{F}}$=50 DRA reduced feature sets selected by rank ordering 1) pre-classification KS-Test p-values and 2) post-classification GRLVQI λ_i relevance values.

For ROC curve generation and analysis, each of the N_D authorized devices presents a *true* claimed ID for itself, as well as, a *false* claimed ID for the other authorized devices (e.g., Dev1 presents a claimed ID for Dev1, Dev2, Dev3, and Dev4). For a specific claimed bit-level ID, N_{TST}=7500 projected *Testing* fingerprints from each of the N_D authorized devices are used to generate (N_{TST}=7500)×(N_D=4)=30000 normalized Multivariate Gaussian posterior probability test statistics according to (3.14). The collection of test statistics are used to create the In-Class and Out-of-Class Probability Mass Functions (PMFs) described in Section 3.5.3 for the specific claimed ID. For example, the In-Class PMF is constructed from 7500 test statistics where the current RF-DNA fingerprint is indeed from the *true* claimed device ID; this same In-Class PMF is subsequently used for *Rogue Device Rejection* assessment in Sect. 4.5.2. The associated Out-of-class PMF is constructed from 22500 test statistics where the current RF-DNA fingerprint is from a *falsely* claimed device ID. Representative PMFs are presented in Fig. 4.7 for one specific case where all N_D=4 authorized devices present claimed bit-level IDs for Dev2. The resultant In-Class and Out-of-Class PMFs are used to produce one *Authorized Device ID Verification* Receiver Operating Characteristics (ROC) curve.

Figure 4.8 shows *Authorized Device ID Verification* performance for each of the N_D=4 authorized ZigBee devices for a DRA reduced feature set of $N_{\widehat{F}}$=50 features selected using 1) pre-classification KS-Test values, and 2) post-classification GRLVQI relevance rankings. The verification ROC curves were generated at SNR=18 dB which corresponds to the $\%C$=90% benchmark in Fig. 4.6 using the same feature set. The N_D=4 ROC curves show that there exists a device-dependent verification threshold t_V(m) such that all authorized



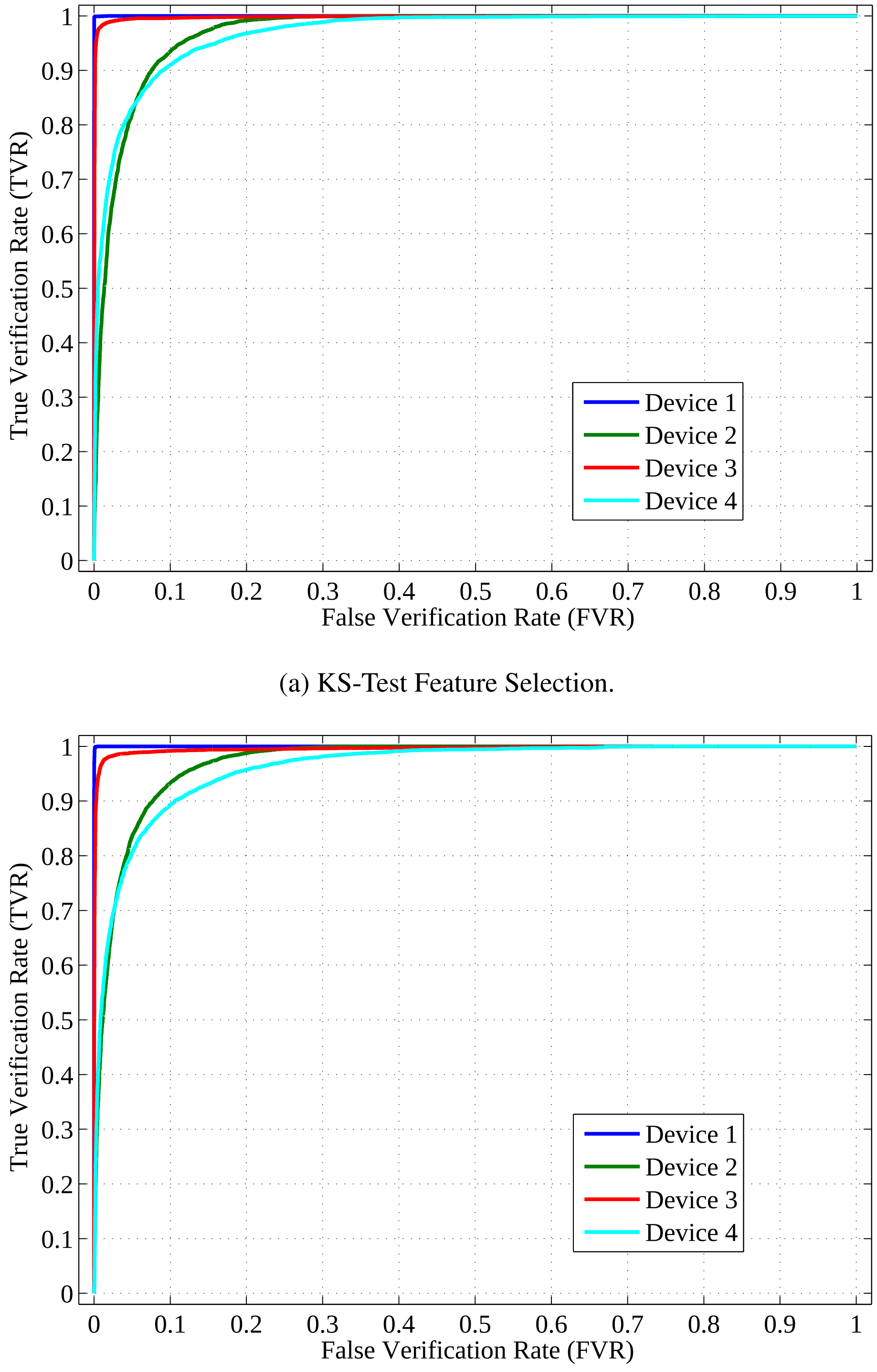
(a) In-Class PMF: Device 2, 7,500 *Testing* RF-DNA fingerprints.



(b) Out-of-Class PMF: Devices (1,3,4), 22,500 *Testing* RF-DNA fingerprints.

Figure 4.7: In-Class and Out-of-Class PMFs for Claimed ID = Device 2. Generated from test statistic z_V in (3.14) for KS-Test top-ranked $N_{\widehat{F}}$=50 features at SNR=18 dB.

device IDs can be verified at True Verification Rate (TVR>90%) and False Verification Rate (FVR<10%) for both methods considered.



(a) KS-Test Feature Selection.

(b) GRLVQI Feature Selection.

Figure 4.8: ZigBee *Authorized Device ID Verification* for N_D=4 authorized devices operating at SNR=18.0 dB (%C≈90% in Fig. 4.6) using top-ranked $N_{\widehat{F}}$=50 features from (a) pre-classification KS-Test and (b) post-classification GRLVQI selection methods.

4.5.2 Rogue Device Rejection.

The ability to use RF-DNA to reject unauthorized rogue devices presenting false bit-level identities is demonstrated using the same ID verification process used for authorized devices. *Rogue Device Rejection* is an assessment of “how well” *current* RF-DNA fingerprints from a pool of rogue (previously unseen and unauthorized) devices match RF-DNA fingerprints associated with the *claimed* ID of an authorized device. This is demonstrated here using N_R=6 (Dev5, Dev6, Dev7, Dev8, Dev9, Dev10) unauthorized rogue devices whose emissions were collected under various conditions (“CAGE”, “LOS”, and “WALL”). A total of N_{TST}=(1000 SHR)×(1 Location)×(5 N_{Nz})=5000 previously unseen RF-DNA fingerprint realizations were used for each of the N_R devices. Table 4.2 lists the 9 ZigBee device ID and collection condition combinations that were considered using the N_R=6 rogue devices. For each of the 9 different combinations, the rogue device presented a claimed ID for each of the N_D=4 authorized device, producing a total of 36 *Rogue Device Rejection* scenarios.

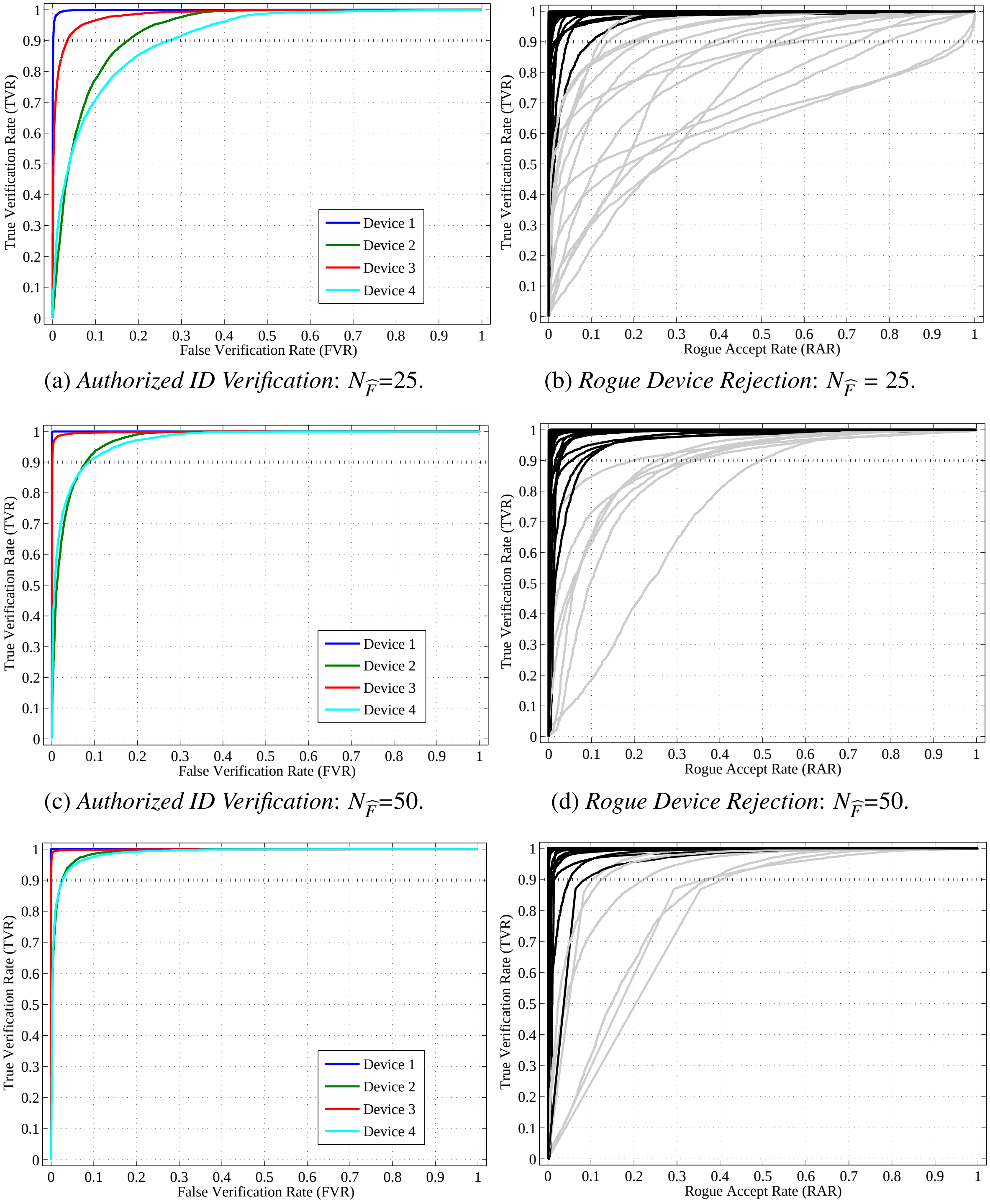
Table 4.2: Nine ZigBee Device ID and collection condition combinations used for Assessing *Rogue Device Rejection* capability. Grey cells correspond untested combinations.

ZigBee ID	CAGE	LOS	WALL
Dev5	▨	X	X
Dev6	▨	X	X
Dev7	▨	X	X
Dev8	X	▨	▨
Dev9	X	▨	▨
Dev10	X	▨	▨

For a specific claimed bit-level ID, N_{TST}=5000 projected *Testing* fingerprints from a rogue device are used to generate 5000 test statistics using (3.14). The collection of test statistics are used to construct the Out-of-Class PMF which is used with the corresponding claimed ID In-Class PMF generated as part of the *Authorized Device ID Verification* process in Sect. 4.5.1. The resultant PMFs are used to produce one ROC performance curve. As detailed in the following two subsections, *Rogue Device Rejection* capability was assessed using each of the DRA feature selection methods.

4.5.2.1 KS-Test Selected Features.

Results for *Rogue Device Rejection* assessment using the KS-Test DRA selected features are presented in Fig. 4.9. These results include the 36 rogue scenarios using top-ranked $N_{\widehat{F}}$=25, 50, 100 feature sets at SNR=18.0 dB. These are conventional ROC curves presented as True Verification Rate (TVR) versus Rogue Accept Rate (RAR), where Rogue Reject Rate is defined as RRR=1−RAR; a higher RAR (lower RRR) reflects greater rogue access and poorer network security performance. *Authorized Device ID Verification* ROC curves are provided alongside the rogue device ID ROC curves to enable identification of the fixed threshold that achieves authorized device TVR>90% and direct mapping to the corresponding RAR (RRR) for each rogue scenario. The solid black curves in Fig. 4.9 (b), (d), and (f) correspond to rogue scenarios that achieve an arbitrary RAR<10% (RRR>90%) performance benchmark when the threshold is fixed such that TVR>90%. As indicated, performance using $N_{\widehat{F}}$=25, 50, 100 KS-Test feature sets achieved the arbitrary RRR>90% benchmark in 21, 29, and 30 out of the 36 rogue scenarios, respectively. Table 4.3 through Table 4.5 highlight rogue scenarios which fail to achieve the arbitrary RRR>90% performance benchmark using selected DRA feature subsets.



(a) *Authorized ID Verification*: $N_{\widehat{F}}$=25.

(b) *Rogue Device Rejection*: $N_{\widehat{F}} = 25$.

(c) *Authorized ID Verification*: $N_{\widehat{F}}$=50.

(d) *Rogue Device Rejection*: $N_{\widehat{F}}$=50.

(e) *Authorized ID Verification*: $N_{\widehat{F}}$=100.

(f) *Rogue Device Rejection*: $N_{\widehat{F}}$=100.

Figure 4.9: Performance using ***KS-Test*** selected features ($N_{\widehat{F}}$=25, 50, 100) for N_D=4 authorized devices and N_R=6 unauthorized rogue devices in various operating scenarios falsely claiming each of the N_D=4 authorized device IDs (36 total rogue scenarios). Grey ROC curves correspond to rogue scenarios where *RAR*<10% (*RRR*>90%) is not achieved.

Table 4.3: ZigBee device ID and collection condition combinations from Table 4.2 where *Rogue Device Rejection* performance in Fig. 4.9 fails to meet *RAR*<10% (*RRR*>90%) with $N_{\widehat{F}}$=25 features selected using KS-Test DRA at SNR=18 dB. The numbers correspond to the Rogue device claimed ID and indicate failure for 15 of 36 rogue scenarios.

ZigBee ID	CAGE	LOS	WALL
Dev5	▨	3	1,3
Dev6	▨	1,3	1,3
Dev7	▨	1,3	1,3
Dev8	1	▨	▨
Dev9	1	▨	▨
Dev10	3,4	▨	▨

Table 4.4: ZigBee device ID and collection condition combinations from Table 4.2 where *Rogue Device Rejection* performance in Fig. 4.9 fails to meet *RAR*<10% (*RRR*>90%) with $N_{\widehat{F}}$=50 features selected using KS-Test DRA at SNR=18 dB. The numbers correspond to the Rogue device claimed ID and indicate failure for 7 of 36 rogue scenarios.

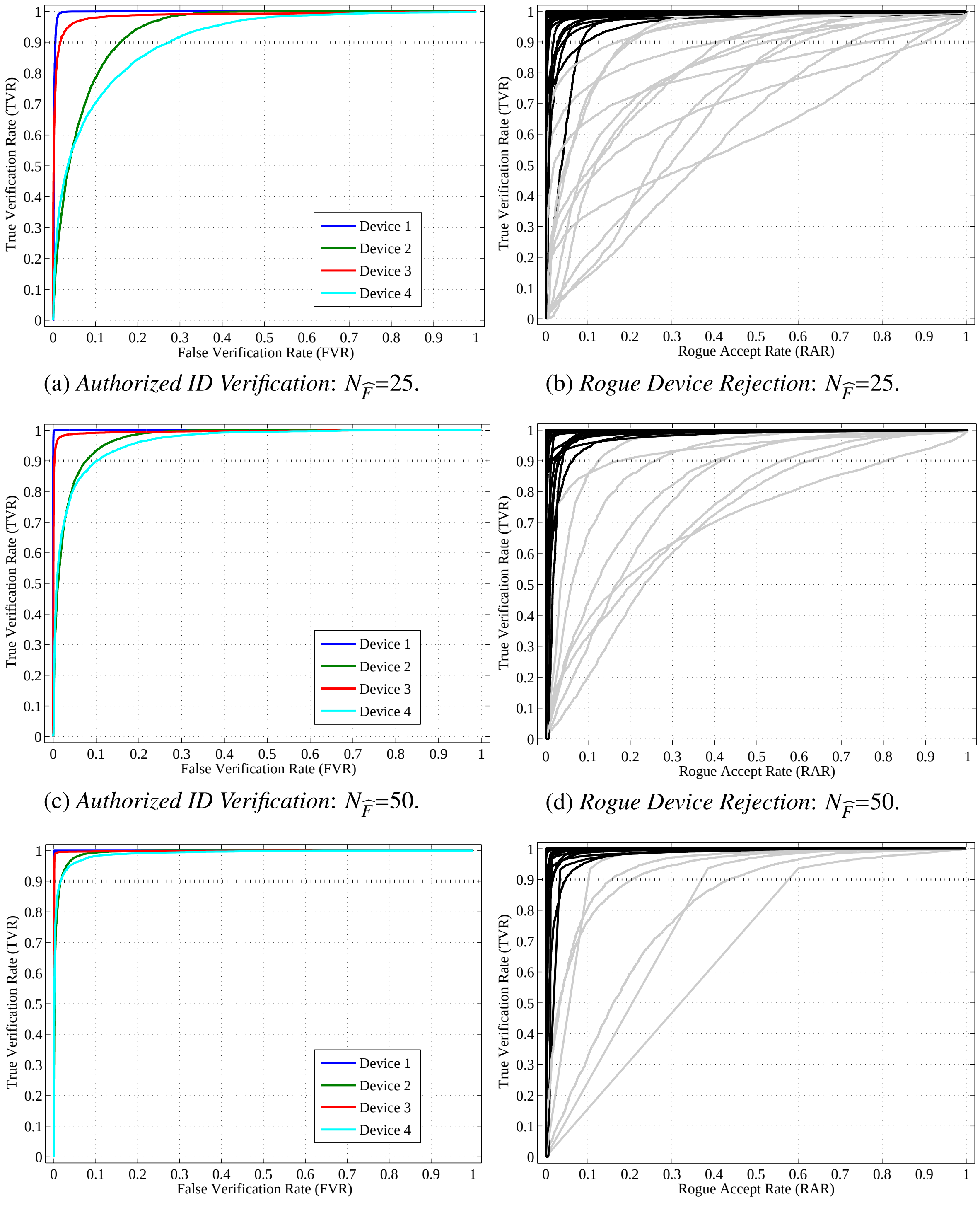
ZigBee ID	CAGE	LOS	WALL
Dev5	▨	3	1
Dev6	▨		3
Dev7	▨		1
Dev8	1	▨	▨
Dev9	1	▨	▨
Dev10	4	▨	▨

Table 4.5: ZigBee device ID and collection condition combinations from Table 4.2 where *Rogue Device Rejection* performance in Fig. 4.9 fails to meet RAR<10% (RRR>90%) with $N_{\widehat{F}}$=100 features selected using KS-Test DRA at SNR=18 dB. The numbers correspond to the Rogue device claimed ID and indicate failure for 6 of 36 rogue scenarios.

ZigBee ID	CAGE	LOS	WALL
Dev5	■	3	1
Dev6	■		3
Dev7	■		1
Dev8		■	■
Dev9	1	■	■
Dev10	4	■	■

4.5.2.2 GRLVQI Selected Features.

Results for *Rogue Device Rejection* assessment using the GRLVQI DRA selected features are presented in Fig. 4.10. These results include the 36 rogue scenarios using top-ranked $N_{\widehat{F}}$=25, 50, 100 feature sets at SNR=18.0 dB. As with KS-Test results presented in Sect. 4.5.2.1, an arbitrary RRR>90% benchmark is used for comparative assessment at an *Authorized Device ID Verification* operating point of TVR>90%. The solid black curves in Fig. 4.10 (b), (d), and (f) correspond to rogue scenarios that achieve the arbitrary RRR>90% benchmark for a fixed threshold yielding TVR>90%. As indicated, performance using $N_{\widehat{F}}$=25, 50, 100 GRLVQI feature sets achieved the arbitrary RRR>90% benchmark in 23, 28, and 30 out of the 36 rogue scenarios, respectively. Table 4.6 through Table 4.8 highlight rogue scenarios which fail to achieve the arbitrary RRR>90% performance benchmark using selected DRA feature subsets.



(a) *Authorized ID Verification*: $N_{\widehat{F}}$=25.

(b) *Rogue Device Rejection*: $N_{\widehat{F}}$=25.

(c) *Authorized ID Verification*: $N_{\widehat{F}}$=50.

(d) *Rogue Device Rejection*: $N_{\widehat{F}}$=50.

(e) *Authorized ID Verification*: $N_{\widehat{F}}$=100.

(f) *Rogue Device Rejection*: $N_{\widehat{F}}$=100.

Figure 4.10: Performance using ***GRLVQI*** selected features ($N_{\widehat{F}}$=25, 50, 100) for N_D=4 authorized devices and N_R=6 unauthorized rogue devices in various operating scenarios falsely claiming each of the N_D=4 authorized device IDs (36 total rogue scenarios). Grey ROC curves correspond to rogue scenarios where *RAR*<10% (*RRR*>90%) is not achieved.

Table 4.6: ZigBee device ID and collection condition combinations from Table 4.2 where *Rogue Device Rejection* performance in Fig. 4.9 fails to meet *RAR*<10% (*RRR*>90%) using $N_{\widehat{F}}$=25 features selected using GRLVQI DRA at SNR=18 dB. The numbers correspond to the Rogue device claimed ID and indicate failure for 13 of 36 rogue scenarios.

ZigBee ID	CAGE	LOS	WALL
Dev5	■	3	1
Dev6	■	3	3,4
Dev7	■	1,3	1,3
Dev8	1	■	■
Dev9	1	■	■
Dev10	3,4	■	■

Table 4.7: ZigBee device ID and collection condition combinations from Table 4.2 where *Rogue Device Rejection* performance in Fig. 4.9 fails to meet *RAR*<10% (*RRR*>90%) with $N_{\widehat{F}}$=50 features selected using GRLVQI DRA at SNR=18 dB. The numbers correspond to the Rogue device claimed ID and indicate failure for 8 of 36 rogue scenarios.

ZigBee ID	CAGE	LOS	WALL
Dev5	■	3	1
Dev6	■	3	3
Dev7	■		1
Dev8		■	■
Dev9	1	■	■
Dev10	3,4	■	■

Table 4.8: ZigBee device ID and collection condition combinations from Table 4.2 where *Rogue Device Rejection* performance in Fig. 4.9 fails to meet RAR<10% (RRR>90%) with $N_{\widehat{F}}$=100 features selected using GRLVQI DRA at SNR=18 dB. The numbers correspond to the Rogue device claimed ID and indicate failure for 6 of 36 rogue scenarios.

ZigBee ID	CAGE	LOS	WALL
Dev5	▨	3	1
Dev6	▨		3
Dev7	▨		1
Dev8		▨	▨
Dev9	1	▨	▨
Dev10	4	▨	▨

V. Summary and Conclusions

THIS chapter provides a summary of research activities, research contributions, and recommendations for further research.

5.1 Summary

This research was conducted to expand AFIT's RF "Distinct Native Attribute" DNA (RF-DNA) fingerprinting process to support IEEE 802.15.4 ZigBee communication system applications. ZigBee-based wireless networks are energy efficiency, low complexity, low cost, and widely deployed in many applications, including energy management and efficiency, home, building, and industrial control automation, and home area networks to name a few [14, 45, 46]. As ZigBee networks continue to increase in popularity, higher levels of security become essential and are critical to protect sensitive personal information and physical system access. The particular security concern addressed under this research is the exploitation of bit-level device identities (ID) to gain unauthorized network access.

To counter bit-level "spoofing" attacks, RF-DNA fingerprints are extracted from Physical (PHY) waveform features and used to achieve human-like discrimination of ZigBee network devices in a typical operational environment. By designating certain devices as *authorized* and others as *unauthorized*, ZigBee network vulnerability to outsider threats is assessed using Receiver Operating Characteristic (ROC) curves to characterize both *Authorized Device ID Verification* performance (granting network access to authorized users presenting *true* bit-level credentials) and *Rogue Device Rejection* performance (denying network access to unauthorized rogue devices presenting *false* bit-level credentials).

For demonstrations here, emissions were collected from TI CC2420 ZigBee devices operating under three environmental scenarios: 1) "CAGE"–devices and collection receiver antenna both in an anechoic chamber, 2) "LOS"–devices within Line-of-Sight of the collection receiver antenna, and 3) "WALL"–devices placed behind a wall relative to the collection receiver antenna. For each device, RF-DNA fingerprint features were extracted from a "hybrid" pool of emissions containing emissions from each of the operational environments. The hybrid features were used for Multiple Discriminant Analysis (MDA) model development and Maximum Likelihood (ML) *Device Classification* performed using both full-dimensional and Dimensional Reduction Analysis (DRA) reduced dimensional RF-DNA fingerprints. The DRA reduced sets were selected using a 1) pre-classification Kolmogorov-Smirnov (KS)-test process and 2) post-classification Generalized Relevance Learning Vector Quantization-Improved (GRLVQI) feature relevance ranking process. The same hybrid MDA/ML model was used in a verification process for assessing *Authorized Device ID Verification* and *Rogue Device Rejection*. In both cases, devices attempt to gain network access by providing bit-level ID credentials (ZigBee MAC address); authorized devices present *true* bit-level IDs while rogue devices present *false* bit-level IDs matching authorized device IDs. The 1 vs. 1 verification process extracts RF-DNA fingerprints from a *current* device emission and compares it with *stored* RF-DNA fingerprint for the claimed ID. Network access is granted (rightly or wrongly) based on a measure of similarity (test statistic) that provides a "Looks how much like?" assessment of the two RF-DNA fingerprints.

5.2 Conclusions

Using device RF-DNA features remains a viable alternative for augmenting bit-level security protocols. This is supported by results here which show that RF-DNA from IEEE 802.15.4 Zigbee emissions can be used as inputs to an MDA/ML discrimination process to perform reliable 1 vs. N_D "Looks most like?" classification assessment, as well as 1 vs. 1 "Looks how much like?" verification assessment. Performance was first assessed with an MDA/ML model developed using features from a "hybrid" pool of emissions from

N_D=4 devices and full-dimensional RF-DNA fingerprints comprised of N_F=729 features. *Device Classification* performance achieved an arbitrary benchmark of average correct classification $\%C$>90% (across all devices) for $SNR \geq 10.0$ dB, with individual devices achieving $\%C$>80% at this same SNR.

The full-dimensional N_F=729 feature set was reduced using DRA and resultant classification and verification performance assessed. The top-ranked $N_{\widehat{F}}$=243 ZigBee feature subset was *qualitatively* selected according to related work in [31], and *quantitatively* selected using two methods, including: 1) pre-classification KS-Test p-value ranking [12, 31], and 2) post-classification GRLVQI λ_i relevance ranking [12, 33, 36]. Hybrid MDA/ML *Device Classification* performance using these DRA≈66% reduced subsets was marginally poorer than full dimensional performance and reflected a "gain" of $G \approx -1.0$ dB at the $\%C$=90% benchmark; gain is defined herein as the reduction in required SNR, expressed in dB, for two systems, methods, etc., to achieve a given $\%C$ performance. Thus, the implementation trade-off is a 66% reduction in the number of features (computational complexity, storage, etc., reduction) at the expense of requiring an additional $SNR \approx 1.0$ dB improvement in channel conditions.

Additional quantitative KS-Test and GRLVQI DRA feature selection was performed and classification performance assessed using the top-ranked $N_{\widehat{F}}$=200, 100, 50, and 25 features. Relative to the %C>90% benchmark [12]:

1. The KS-Test selected feature sets required $SNR \approx 10.0$ dB ($N_{\widehat{F}}$=243) to $SNR \approx 18.0$ dB ($N_{\widehat{F}}$=50), with results for $N_{\widehat{F}}$=25 failing to meet the benchmark.

2. The GRLVQI selected feature sets required the same $SNR \approx 10.0$ dB ($N_{\widehat{F}}$=243) to $SNR \approx 18.0$ dB ($N_{\widehat{F}}$=50), with results for $N_{\widehat{F}}$=25 achieving the benchmark at $SNR \approx 30.0$ dB.

Hybrid MDA/ML verification performance was assessed for 1) N_D=4 authorized network devices and 2) N_R=6 unauthorized (rogue) network devices. Performance was evaluated using the $N_{\widehat{F}}$=50 DRA feature set at SNR=18.0 dB given that the %C=90% benchmark was achieved under these conditions. ROC curve analysis for *Authorized Device ID Verification* indicated that there exists a device dependent threshold $t_V(m)$ for *all* authorized devices such that a True Verification Rate of TVR>90% and False Verification Rate of FVR<10% are realized for both DRA methods; this range of TVR and FVR was arbitrarily selected for comparative assessment.

Rogue Device Rejection capability was assessed using N_R=6 unauthorized devices placed in nine collection combinations of various experimental "CAGE", "LOS", and "WALL" locations, with each rogue device falsely presenting a claimed ID matching each of the N_D=4 authorized IDs; a total of 36 rogue assessment scenarios. Considering an arbitrary Rogue Rejection Rate of *RRR*>90%, ROC curve analysis for *Rogue Device Rejection* indicated that performance using KS-Test and GRLVQI selected feature sets was consistent. Specific performance included [12]:

1. The KS-test selected feature sets achieving *RRR*>90% in 21, 29, and 30 of the 36 rogue scenarios using $N_{\widehat{F}}$=100, 50, and 25 top-ranked features, respectively.
2. The GRLVQI selected feature sets achieving *RRR*>90% in 23, 28, and 30 of the 36 rogue scenarios using $N_{\widehat{F}}$=100, 50, and 25 top-ranked features, respectively.

5.3 Recommendations for Future Research

This research provides a proof-of-concept demonstration that highlights the promise for augmenting ZigBee bit-level security mechanisms. This was done using RF-DNA features with an MDA/ML discrimination process. The work here is by no means complete and there are several potential directions that future research could take:

1. Performing a detailed assessment of ZigBee GRLVQI DRA–Results here for dimensionally reduced feature sets were based on two separate rank-ordering and selection methods (pre-classification KS-Test and post-classification GRLVQI) being

developed in parallel under AFIT's RF Intelligence (RFINT) program. GRLVQI parameter settings and model development were *not* optimized for ZigBee emissions as part of this research. Further analysis and GRLVQI optimization could be done to better exploit feature set dependence, or independence, as collection location varies ("CAGE", "LOS", "WALL") and environmental conditions change.

2. Increasing the number of model training devices–An iterative process should be considered for progressively expanding the pool of authorized devices being used for model development. The less than perfect *Rogue Device Rejection* performance here ($RRR \neq 100\%$) was not too surprising given that 1) MDA model development is a classification-based versus verification-based optimization process and similar results have been observed using other signals, and 2) only N_D=4 authorized devices were used for hybrid MDA/ML model development; it is highly unlikely that RF-DNA features from N_D=4 population members of a larger population (thousands or even millions) accurately capture population behavior and provide broad human-like discrimination. Increasing the sample size (training devices) will allow the developed models to better represent the larger device population.

3. Considering alternate test statistics–Results here are based exclusively on inherent MATLAB functionality for implementing MDA model development and performing ML classification assessment (*classify* function), as well as, ROC curve (*roc* function) verification performance assessment; the inherent normalized MVG posterior probability similarity measure was used exclusively as the test statistic. There are a myriad of additional probability-based, as well as distance-based, similarity measures that could be considered and which may improve overall performance.

Bibliography

[1] "OSI Reference Model", May 2009. URL http://www.ccnaguru.com/osi-reference-model.html.

[2] Agilent Technologies Inc., USA. *Agilent E3238 Signal Intercept and Collection Solutions: Family Overview*, Publication 5989-1274EN, Jul 2004.

[3] Agilent Technologies Inc., USA. *E3238S Command Reference*, Publication E3238-90012, 2009.

[4] Buckner M.A. *Learning From Data with Localized Regression and Differential Evolution*. Ph.D. thesis, University of Tennessee, Knoxville, May 2003.

[5] Buckner M.A., A.M. Urmanov, A.V. Gribok and J.W. Hines. "Application of Localized Regularization Methods for Nuclear Power Plant Sensor Calibration Monitoring", Technical Correspondence, 2002.

[6] Buckner M.A., M. Bobrek, E.E. Farquahar, Harmer P.K. and M.A. Temple. "Enhancing Network Security Using 'Learning-From-Signals' and Fractioinal Fourier Transform Based RF-DNA Fingerprints". *SDR'11- Wireless Innovation Conference*. Dec 2011.

[7] Chen, S.K., T. Kao, C.T. Chan, C.N. Huang, C.Y. Chiang, C.Y. Lai, T.H. Tung, and P.C. Wang. "A Reliable Transmission Protocol for ZigBee-Based Wireless Patient Monitoring". *IEEE Trans on Information technology in Biomedicine*, 16(1):6–16, Jan 2012.

[8] Danev B. and S. Kapkun. "Transient-Based Identification of Wireless Sensor Nodes". *Proc of the 8th ACM/IEEE Int'l Conf on Information Processing in Sensor Networks (IPSN09)*. Apr 2009.

[9] Danev B., H. Luecken, S. Capkun, and K. El Defrawy. "Attacks on Physical-layer Identification". *Proc of the 3rd ACM Int'l Conf on Wireless Network Security (WiSec10)*. Mar 2010.

[10] Danev B., Heydt-Benjamin, S. Thomas and S. Capkun. "Physical-layer Identification of RFID Devices". *Proc of the 18th conference on USENIX security symposium*, SSYM'09, 199–214. 2009.

[11] Dubendorfer C.K., B.W. Ramsey and M.A. Temple. "An RF-DNA Verification Process for ZigBee Networks". *Proc of Military Communications Conference (MILCOM12)*. Oct 2012.

[12] Dubendorfer C.K., B.W. Ramsey and M.A. Temple. “ZigBee PHY-Based Device ID Verification: Security Enhancement for Industrial Control and Building Automation Systems”. *Proc of IFIP Working Group 11.10 Int’l Conf on Critical Infrastructure Protection (IFIP13)*. Mar 2013.

[13] Duda R., P. Hart and D. Stork. *Pattern Classification*. John Wiley & Sons, Inc., New York, second edition, 2001.

[14] Egan D. “The Emergence of ZigBee in Building Automation and Industrial Control”. *Computing & Control Engineering Journal*, 16(2):14–19, April-May 2005.

[15] Fawcett T. *ROC Graphs: Notes and Practical Considerations for Researchers*. Kluwer Academic Publishers, Netherlands, Mar 2004.

[16] Hall J., et. al. “Detection of Transient in Radio Frequency Fingerprinting Using Signal Phase”. IASTED Int’l Conf on Wireless and Optical Communications, May 2003.

[17] Hall J., et. al. “Using Transceiverprints for Anomaly Based Intrusion Detection”. 3rd IASTED Int’l Conf on Communications, Internet and Information Technology (CIIT), November 2004.

[18] Hall J., M. Barbeau and E. Kranakis. “Radio Frequency Fingerprinting for Intrusion Detection in Wireless Networks”, Jul 2005. DRAFT.

[19] Hall J., M. Barbeau and E. Kranakis. “Detecting Rogue Devices In Bluetooth Networks Using Radio Frequency Fingerprinting”. *Communications and Computer Networks*, 108–113. 2006.

[20] Harmer P.K. *Development of Learning from Signals Classifier for Cognitive Software Defined Radio Applications*. Ph.D. thesis, Air Force Institute of Technology, March 2013.

[21] Harmer P.K., D.R. Reising and M.A. Temple. “Classifier Selection for Physical Layer Security Augmentation in Cognitive Radio Networks”. Proc of IEEE Int’l Conf on Communications (ICC13), Jun 2013.

[22] Harmer P.K., M.A. Temple, M.A. Buckner and E.E. Farquahar. “4G Security Using Physical Layer RF-DNA with DE-Optimized LFS Classificatioin”. *Jour of Communications, Special Issue: Advances in Communications and Networking*, 9(6):671–681, Dec 2011.

[23] Harmer P.K., M.A. Temple, M.A. Buckner and E.E. Farquahar. “Using Differential Evolution to Optimize ’Learning from Signals’ and Enhance Network Security”. *Submitted to: Genetic and Evolutionary Computation Conference (GECCO)*. July Jul 2011.

[24] Harmer P.K., M.D. Williams and M.A. Temple. "Using DE-Optimized LFS Processing to Enhance 4G Communication Security". *20th International Conference on Computer Communication and Networks (ICCCN)*. Aug 2011.

[25] Hastie T., R. Tibshirani and J. Friedman. *The Elements of Statistical Learning; Data Mining, Inference, and Prediction*. Springer-Verlag, New York, New York, USA, 2001. ISBN 0-387-95284-5.

[26] Institute of Electrical and Electronics Engineers. *IEEE 802.15.4, Standard, Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate WPANS*, 2006.

[27] Jana S. and S.K. Kasera. "Wireless Device Identification with Radiometric Signatures". *Proc of the ACM 14th Int'l Conf on Mobile Computing and Networking (MOBICOM08)*. Sep 2008.

[28] Klein R.W. *Application of Dual-Tree Complex Wavelet Transforms to Burst Detection and RF Fingerprint Classification*. Ph.D. thesis, Air Force Institute of Technology, September 2009.

[29] Klein R.W., M.A. Temple and M.J. Mendenhall. "Application of Wavelet-Based RF Fingerprinting to Enhance Wireless Network Security". *Jour of Communications and Networks*, Vol. 11, No. 6, Dec 2009.

[30] Klein R.W., M.A. Temple, M.J. Mendenhall and D.R. Reising. "Sensitivity Analysis of Burst Detection and RF Fingerprinting Classification Performance". *Proc of IEEE Int'l Conf on Communications (ICC09)*. Jun 2009.

[31] Ramsey B.W., M.A. Temple and B.E. Mullins. "PHY Foundation for Multi-Factor ZigBee Node Authentication". *Proc of IEEE Global Communications Conf (GLOBECOM12)*. Dec 2012.

[32] Ramsey Electronics Inc. *The 'Logi' Log Periodic Antenna: Ramsey Electronics Model No. LPY2*. URL http://www.ramseyelectronics.com/downloads/manuals/LPY2.pdf.

[33] Reising, D.R. *Exploitation of RF-DNA for Device Classification and Verification Using GRLVQI Processing*. Ph.D. thesis, Air Force Institute of Technology, December 2012.

[34] Reising D.R. and M.A. Temple. "WiMAX Mobile Subscriber Verification Using Gabor-Based RF-DNA Fingerprints". *2012 IEEE International Communications Conference (ICC12)*. Jun 2012.

[35] Reising, D.R., and M.A. Temple. "Verification of Localized OFDM-Based Devices Using Dimensionally Efficient GRLVQI Processing". *Systems Journal, IEEE*, XX(##):xx–xx, 2012, UNDER REVIEW.

[36] Reising D.R., M.A. Temple and J.A. Jackson. "Dimensional Reduction Analysis Using RF Fingerprints in Support of Enhanced WSN-Based Industrial Control System Security". *Security and Communication Networks Journal (SCN-SI-039)*, UNDER REVIEW 2012.

[37] Reising D.R., M.A. Temple, and J.A. Jackson. "Detecting Rogue Devices at Cloud Wireless Access Points Using RF Air Monitors". *Information Forensics and Security, IEEE Transactions on*, XX(##):xx–xx, 2012, UNDER REVIEW.

[38] Reising D.R., M.A. Temple and M.E. Oxley. "Gabor-Based RF-DNA Fingerprinting for Classifying 802.16e WiMAX Mobile Subscribers". 2012 IEEE Int'l Conf on Computing, Networking & Communications (ICNC), Jan 2012.

[39] Reising D.R., M.A. Temple and M.J. Mendenhall. "Improved Wireless Security for GMSK-Based Devices Using RF Fingerprinting". *Int. J. Electronic Security and Digital Forensics*, Vol. 3, No. 1, pp. 41-59,2010.

[40] Reising D.R., M.A. Temple and M.J. Mendenhall. "Improving Intra-Cellular Security Using Air Monitoring with RF Fingerprints". *Proc of 2010 IEEE Wireless Communications & Networking Conf (WCNC10)*. Apr 2010.

[41] Speers R., J. Wright and R. Melgares. "Api-do: Tools for ZigBee and 802.15.4 Security Auditing". URL http://code.google.com/p/zigbeesecurity/.

[42] Suski W.M. II, M.A. Temple, M.J. Mendenhall and R.F. Mills. "RF Fingerprinting Commercial Communication Devices to Enhance Electronic Security". *Int. J. Electronic Security and Digital Forensics*, Vol. 1, No. 3, pp. 301-322, 2008.

[43] Suski W.M. II, M.A. Temple, M.J. Mendenhall and R.F. Mills. "Using Spectral Fingerprints to Improve Wireless Network Security". *Proc of IEEE Global Communications Conf (GLOBECOM08)*. Mar 2008.

[44] Theodoridis S. and K. Koutoumbas. *Pattern Recognition*. Academic Press, fourth edition, 2009.

[45] Tihon I. and V. Croitoru. "ZigBee Sensor Networks Telesurveillance". 1–4. 10th International Symposium on Signals, Circuits and Systems (ISSCS'11), Jun 2011.

[46] Whittaker T. "Final Word". 18(3):48, Jun-July 2007.

[47] Williams M.D., M.A. Temple and D.R. Reising. "Augmenting Bit-Level network Security Using Physical Layer RF DNA Fingerprinting". *Proc of IEEE Global Communications Conf (GLOBECOM10)*. Dec 2010.

[48] Williams M.D., S.A. Munns, M.A. Temple and M.J. Mendenhall. "RF-DNA Fingerprinting for Airport WiMax Communications Security". *Proc of 4th Int'l Conf on Net and Sys Security (NSS10)*. Sep 2010.

[49] Wright J. “KillerBee: Framework and Tools for Exploiting ZigBee and IEEE 802.15.4 Networks”, Version 1.0, 2010. URL http://code.google.com/p/killerbee/.

[50] Yang H. and S.H. Yang. “Connectionless Indoor Inventory Tracking in ZigBee RFID Sensor Network”. 2618–2623. 25th Annual Industrial Electronics Conference (IECON’09), Nov 2009.

[51] ZigBee Alliance. *ZigBee Specification*, 2008.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704–0188*

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
21–03–2013	Master's Thesis	Oct 2011–Mar 2013

4. TITLE AND SUBTITLE

Using RF-DNA Fingerprints to Discriminate ZigBee Devices in an Operational Environment

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dubendorfer, Clay K., Civilian, USAF

5d. PROJECT NUMBER

JON# 11G186, 12G186

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB, OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENG-13-M-15

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory, AFMC
Attn: AFRL/RYWE (Dr. Vasu Chakravarthy)
2241 Avionics Circle, Bldg. 620
WPAFB, OH 45433-7734
(937)-798-8269
Vasu.Chakravarthy@wpafb.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/RYWE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This research was performed to expand AFITs Radio Frequency "Distinct Native Attribute" (RF-DNA) fingerprinting process to support IEEE 802.15.4 ZigBee communication network applications. Current ZigBee bit-level security measures include use of network keys and MAC lists which can be subverted through interception and spoofing using open-source hacking tools. This work addresses device discrimination using Physical (PHY) waveform alternatives to augment existing bit-level security mechanisms. ZigBee network vulnerability to outsider threats was assessed using Receiver Operating Characteristic (ROC) curves to characterize both *Authorized Device ID Verification* performance (granting network access to authorized users presenting *true* bit-level credentials) and *Rogue Device Rejection* performance (denying network access to unauthorized rogue devices presenting *false* bit-level credentials). Radio Frequency 'Distinct Native Attribute' (RF-DNA) features are extracted from time-domain waveform responses of 2.4 GHz CC2420 ZigBee transceivers to enable human-like device discrimination. The fingerprints were constructed using a "hybrid" pool of emissions collected under a range of conditions, including anechoic chamber and an indoor office environment where dynamic multi-path and signal degradation factors were present. The RF-DNA fingerprints were input to a Multiple Discriminant Analysis, Maximum Likelihood (MDA/ML) discrimination process and a 1 vs. many "Looks most like?" classification assessment made. The hybrid MDA model was also used for 1 vs. 1 "Looks how much like?" verification assessment. ZigBee *Device Classification* performance was assessed using both full and reduced dimensional fingerprint sets. Reduced dimensional subsets were selected using Dimensional Reduction Analysis (DRA) by rank ordering 1) pre-classification KS-Test p-values and 2) post-classification GRLVQI λ_i feature relevance values. Assessment of Zigbee device ID verification capability included both *Authorized Device ID Verification* and *Rogue Device Rejection*.

15. SUBJECT TERMS

RF-DNA, Fingerprinting, ZigBee, Classification, Verification, Network Security

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Dr. Michael A. Temple (ENG)
					19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	86	(937) 255-3636 x4279 Michael.Temple@afit.edu